

Northeast Decision Sciences Institute

54th Annual Conference

March 27-29, 2025

Hershey, PA

*Generative AI to Enhance
Supply Chain Resilience*



Decision Sciences Institute

Scholars • Educators • Problem Solvers

NEDSI 2025

The board of the Northeast Decision Sciences Institute gratefully acknowledges the host institution for our 54th annual conference:

**Penn State Harrisburg
School of Business Administration**



PennState
Harrisburg

March 27-29, 2025

Hershey, PA





WE ARE

PENN STATE in the Capital Region



PennState
Harrisburg

School of
Business Administration

717-948-6142 • amy.atkins@psu.edu

harrisburg.psu.edu/sba



ASSOCIATE DEGREE

Business Administration

BACHELOR'S DEGREES

Accounting
Cybersecurity Analytics and Operations
Enterprise Technology Integration
Finance
Human Capital Management
Human-Centered Design and Development
Information Systems
Management
Marketing
Project and Supply Chain Management
Security and Risk Analysis

MASTER'S DEGREES

Accounting
Business Administration
Information Systems

ACCELERATED UNDERGRADUATE/ GRADUATE DEGREES

Accounting/M.B.A.
Finance/M.B.A.
Information Systems/M.B.A.
Information Systems/M.S.I.S.
Management/M.B.A.
Marketing/M.B.A.
Project and Supply Chain Management/M.B.A.



KEAN

College of
**Business and
Public Management**



A rapidly changing world needs bold leaders in business, public management and criminal justice. Kean's College of Business and Public Management equips you with the knowledge, skills and connections to succeed.

Whether you plan to work on Wall Street, lead a company, launch a startup, or serve in government or law enforcement, our programs provide hands-on learning, internships and mentorships that position you for success.

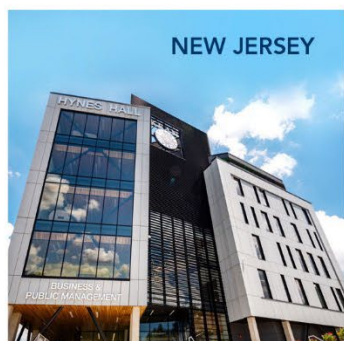
At Kean, you gain more than a degree—you build the expertise and network to lead and make an impact.

UNDERGRADUATE PROGRAMS

- Accounting (B.S.)
- Criminal Justice (B.A.)
- Economics (B.S.)
- Finance (B.S.)
- Global Business (B.S.)
- Management (B.S.)
- Marketing (B.S.)
- Public Administration (B.A. and B.A./MPA)

GRADUATE PROGRAMS

- Accounting (M.S.)
- Business Administration (MBA)
- Criminal Justice (M.A.)
- Public Administration (MPA)



**HOME OF KEAN'S COLLEGE OF
BUSINESS AND PUBLIC MANAGEMENT**

SILVER SPONSOR

EQUIPPING THE LEADERS

FOR EMERGING TECHNOLOGIES

James Madison University College of Business
offering degree programs, minors, concentrations and a non-credit certificate course in areas such as:

- Accounting
- Business Management
- Computer Information Systems
- Economics
- Finance
- International Business
- Marketing
- Quantitative Finance
- Business Analytics
- Entrepreneurship
- Global Supply Chain Management
- Information and Cybersecurity Management
- Information Systems Concentration
- Supply Chain Analytics
- Innovation and Entrepreneurship
- Human Resource Management



LEARN MORE AT:
jmu.edu/cob

JMU
College of Business



■ TABLE OF CONTENTS

Welcome Letters

Dave Callejo Pérez, Chancellor and Dean, Penn State Harrisburg

John Haddad, Interim School Director, Penn State Harrisburg, School of
Business Administration

Dinesh Pai, NEDSI 2025 Program Chair

NEDSI 2025 Keynote Speaker

About the Decision Sciences Institute

About the Northeast Decision Sciences Institute

NEDSI 2025 Conference Sponsors

NEDSI 2025 Lifetime Service Award

NEDSI Board of Directors 2024-2025

Officers at Large 2024-2025

Presidential Advisory Council 2024-2025

Advisors at Large 2024-2025

Program Track Chairs

Conference Reviewers

NEDSI Service Recognition to Past Program Chairs

Service Recognition to Past Presidents

Best Paper Awards Committee

Undergraduate/Master Student Poster Competition Committee

NEDSI 2025 Best Paper Awards

Undergraduate/Graduate Research Poster Participants

Conference Program Schedule of Presentations

Call for Papers for NEDSI 2026 in Philadelphia, PA

Thank You to Our Sponsors

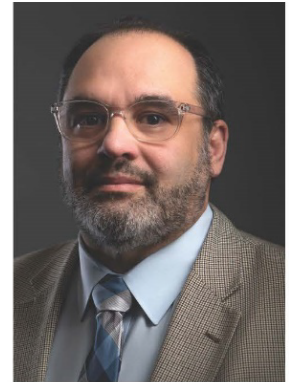
■ Welcome



PennState
Harrisburg

Office of the Chancellor
The Pennsylvania State University
C119 Olmstead Building
777 West Harrisburg Pike
Middletown, PA 17057

717-948-6105
pshchancellor@psu.edu
hbg.psu.edu



March 27, 2025

Dear NEDSI Members and Guests,

As the proud host institution of the annual meeting of the Northeast Decision Sciences Institute, I am honored to welcome you to the 54th meeting of this prestigious conference. Dr. Dinesh Pai, associate professor of supply chain management in the Penn State Harrisburg School of Business, and his team have put together an excellent program and we are pleased to be able to present it to you.

The theme of this year's conference is "Generative AI to Enhance Supply Chain Resilience". I hope you enjoy your interactions with our Penn State Harrisburg faculty and staff and the various program tracks, keynote speakers and panel discussions.

On behalf of the faculty, staff, and students at Penn State Harrisburg, I wish you a productive and successful meeting.

Best regards,

Dave Callejo Pérez, Ed.D.
Chancellor and Dean

■ Welcome



School of Business Administration
777 West Harrisburg Pike
Middletown, PA 17057

Tel. 717-948-6151
Fax 717-948-6456
harrisburg.psu.edu/sba



March 27, 2025

Dear NEDSI Members and Guests,

It is my great pleasure to welcome you to the annual meeting of the Northeast Decision Sciences Institute. We have a tremendous conference in store for you this year.

As all attendees know, academic conferences are not easy to organize. There are thousands of details one must tend to. Over the past several months, I have witnessed the tremendous investment of time and energy that Dr. Dinesh Pai and his team have put into this event. Thanks to their steady work and careful planning, we are now poised to enjoy what will surely be a tremendous experience.

The theme of this year's conference could not be more timely. Like you, I am looking forward to learning what the latest thinking is on generative AI and its impact on supply chains. However, as fascinating as our panels and speakers promise to be, I also invite you to take a few hours to venture out into Hershey. This town has a fascinating history, which you can access at the Hershey Story (guests of Hershey Lodge and The Hotel Hershey receive complimentary tickets to The Hershey Story and Hershey Gardens). So please enjoy exploring Hershey and do not hesitate to ask any of us for suggestions.

On behalf of Penn State Harrisburg's School of Business Administration, I wish you a productive, stimulating, but also fun conference experience.

Sincerely,

A handwritten signature in cursive script that reads "John R. Haddad".

John R. Haddad

■ Welcome



Dinesh R. Pai
Penn State Harrisburg

Welcome to NEDSI 2025!

On behalf of the board and program committee of the Northeast Decision Sciences Institute, I would like to extend a hearty welcome to the 54th Annual Conference in Hershey, PA!

The theme of the conference is “Generative AI to Enhance Supply Chain Resilience.” The program committee has planned an amazing conference around the theme, complete with a range of pertinent multidisciplinary tracks, plenary talks, workshops, panel discussions, and social hours. The conference’s theme coincidentally aligns with the challenges posed by the recent pandemic and the geopolitical conflicts that have an impact on supply chains. It also addresses the recent rise of artificial intelligence, a source of a national debate. As such, over three days, the conference aims to discuss, deliberate, and provide actionable solutions to overcome many of the challenges.

We are thrilled to once again provide you with a diverse range of paper presentations, panel discussions, and workshops, thanks to the many efforts of academics, leaders, and volunteers. A student poster session will also be held. In short, there will be plenty of opportunities for networking and professional growth at the conference.

The board and organizing committee worked very hard to put together a great 2025 conference. Apart from the NEDSI board, we extend our heartfelt appreciation to Penn State Harrisburg, which will serve as institutional host of NEDSI 2025. Specifically, I would like to thank Dr. David M. Callejo Pérez, Chancellor, Penn State Harrisburg, and Dr. John Haddad, the interim director of the School of Business Administration, without whose support the conference could not have taken place.

In closing, the NEDSI board welcomes you to NEDSI 2025. We are confident it will be a valuable, rewarding, and pleasurable conference for you.

Best regards,



Dinesh R. Pai
NEDSI 2025 Conference Chair

Gen AI and Supply Chain Resilience: Future-Proofing Through Innovation

Adriana Souza



Adriana Souza is a data science expert focused on leveraging AI to drive innovation and solve complex business problems. As a Senior Finance Manager at D&H Distributing, she collaborates with supply chain professionals to develop AI-driven solutions for forecasting and pricing optimization. Her expertise enables firms to enhance decision-making and adaptability in an ever-evolving landscape. At this year's NEDSI conference, Adriana will discuss the transformative effects of generative AI on creating resilient and flexible supply chains, while also sharing practical insights from her experience at the intersection of data science and distribution. She holds a bachelor's degree in economics and a master's degree in data science from New College of Florida, along with an MBA from Louisiana State University.

NJIT MAKES LEADERS

MARTIN TUCHMAN SCHOOL OF MANAGEMENT

Forging business innovation with top-tier research.

NJIT

New Jersey Institute
of Technology

UNIVERSITY HEIGHTS • NEWARK, NJ 07102 • NJIT.EDU



■ ABOUT

The Decision Sciences Institute

The Decision Sciences Institute (DSI) is a professional organization of academicians and practitioners interested in the application of quantitative and behavioral methods to the problems of society.

Through national, international and regional conferences, competitions and publications, the DSI provides an international forum for presenting and sharing researching the study of decision processes across disciplines. The DSI also plays a vital role in the academic community by offering professional development activities and job placement services.

Three divisions— Americas, Asia and Pacific, and Europe - operate within the DSI. The Americas division is comprised of five chapters, including NEDSI, and the Asia and Pacific division has three separate chapters. Each division has a representative on the Board of Directors. Chapters have their own elected board of directors and hold annual meetings.

The DSI, an independent nonprofit educational organization, is located in Houston, Texas, where it receives extensive support from the C.T. Bauer College of Business at the University of Houston.

Northeast Decision Sciences Institute

The Northeast Decision Sciences Institute (NEDSI) is one of five regions in the Americas Division of the professional society, the Decision Sciences Institute. This region encompasses the Northeastern United States. NEDSI holds an annual regional meeting each spring that features presentations of original research papers, panels, workshops, and other interesting innovations in the fields of accounting, business analytics, business education, finance, governance, healthcare management, marketing, operations management/operations research, strategic management, supply chain management, sustainability, and technology management, among others. Awards for “Best Paper” in several categories are given each year.

Follow us on Facebook



■ **NEDSI 2025 CONFERENCE SPONSORS**

Penn State Harrisburg

School of Business Administration

Grand Platinum Sponsor

University of Massachusetts Lowell

Manning School of Business

Platinum Sponsor

Goodwill Keystone Area

Gold Sponsor

Kean University

Gold sponsor

New Jersey Institute of Technology

Martin Tuchman School of Management

Gold Sponsor

James Madison University

College of Business

Silver Sponsor

University of Massachusetts Dartmouth

School of Business Administration

Bronze Sponsor

University of the District of Columbia

School of Business and Public Administration

Bronze Sponsor

Penn State Great Valley

School of Business Administration

Bronze Sponsor

■ NEDSI 2025 Lifetime Service Award



Jennifer Swanson, Stonehill College

Professor Jennifer Swanson has been an active participant in the Northeast Decisions Sciences Institute for almost 30 years, serving as President, Program Chair, Board member, Advisory Council member, Archivist, Poster Committee Chair, and other roles. Her PhD is in Management from the University of Rhode Island and she is currently the Chair of the Global Business, Operations & Strategy Department at Stonehill College. In addition, she has spent numerous summers teaching and training at the undergraduate, graduate, and executive level in Southeast Asia. She has published in a wide variety of journals, served on the Board for MicroLoan Foundation USA, participated in a China Fulbright-Hays Group Project, was Co-Director for a U.S. Department of Education grant on Canada, Ireland, and Portugal, participated in a COIL program team teaching U.S. and Bolivian students, and developed travel learning communities involving student travel to Europe, Africa, Central America, and Latin America.

■ NEDSI BOARD OF DIRECTORS 2024-2025

President	Amit Arora , University of the District of Columbia
Immediate Past President	Theologos Homer Bonitsis , New Jersey Institute of Technology
Co-President-Elect	Neil Desnoyers , Saint Joseph's University
Co-President-Elect	Xiangrong Liu , Bridgewater State University
Program Chair 2025	Dinesh Pai , Penn State Harrisburg
Program Administrator 2025	Joy Field , Boston College
Program Chair-Elect 2026	Afrooz Moatari-Kazerouni , Widener University
V.P. Member Engagement	Ju Myung (JM) Song , University of Massachusetts Lowell
V.P. Promotional Activities	Soheil Sibdari , University of Massachusetts Dartmouth
V.P. Technology	EunSu Lee , New Jersey City University
Treasurer	Haleh Karimi , Bellarmine University
Secretary	Raktim Pal , James Madison University
Archivist	Jennifer A. Swanson , Stonehill College

■ OFFICERS AT LARGE 2024 - 2025

Oya Tukul, New Jersey Institute of Technology
Linda Boardman Liu, Boston College
Douglas Hales, University of Rhode Island

■ PRESIDENTIAL ADVISORY COUNCIL 2024-2025

Theologos Homer Bonitsis, New Jersey Institute of Technology
Dinesh Pai, Penn State Harrisburg
Gang Li, Bentley University
Effie Stavoulaki, Bentley University
Eric Stein, Penn State Great Valley
Douglas Hales, University of Rhode Island
Neset Hikmet, University of South Carolina

■ ADVISORS AT LARGE 2024-2025

Joy Field, Boston College
Fariborz Partovi, Drexel University

■ PROGRAM TRACK CHAIRS

Accounting, Finance, and Economics

Theologos Homer Bonitsis, New Jersey Institute of Technology | theologos.bonitsis@njit.edu

Big Data and Business Analytics

EunSu Lee, New Jersey City University | elee3@njcu.edu

Cyber Security, IT and Emerging Technologies

Douglas Hales | dhales@uri.edu

Decision Support Systems (AI, ML, etc.) and Knowledge Management

Jin Fang, Clark University | jinfang@clarku.edu
Vasileios Pavlopoulos, University of Alabama in Huntsville | vasileios.pavlopoulos@uah.edu

Innovative Education, Curriculum and Cases

Ju Myung (JM) Song, University of Massachusetts Lowell | JuMyung_Song@uml.edu

Healthcare and Service Management

Fatma Pakdil, Eastern Connecticut State University | pakdiltf@easternct.edu

Innovation and Creativity

Neil Desnoyers, St. Joseph's University | ndesnoye@sju.edu

Legal, Ethical and Social Issues

Alisha Harper, Bellarmine University | aharper@bellarmine.edu

Marketing and Consumer Behavior

Anshu Arora, University of the District of Columbia | anshu.arora@udc.edu

Operations Management/Operations Research

Gang Li, Bentley University | gli@bentley.edu

Public Administration and Policy

Anshu Arora, University of the District of Columbia | anshu.arora@udc.edu

Public Administration and Policy

Amit Arora, University of the District of Columbia | amit.arora@udc.edu

Strategy and Organizational Behavior

Haleh Karimi, Bellarmine University | hkarimi@bellarmine.edu

Supply Chain and Logistics Management

Raktim Pal, James Madison University | palrx@jmu.edu

Sustainability Management

Xiagrong Liu, Bridgewater State University | xliu@bridgew.edu

Undergraduate/Master Student Posters

Jennifer Swanson, Stonehill College | jswanson@stonehill.edu
Kellyann Berube Kowalski, University of Massachusetts Dartmouth | kkowalski@umassd.edu
Kathleen Ferris-Costa, Bridgewater State University | kferriscosta@bridgew.edu
Krista Hill, Babson College | khill@babson.edu

■ CONFERENCE REVIEWERS

Amit Arora	University of the District of Columbia
Anshu Arora	University of the District of Columbia
Theologos Homer Bonitsis	New Jersey Institute of Technology
Melike Hazal Can	Harvard University
Jonathan Daigle	Monmouth University
Neil Desnoyers	Saint Joseph's University
Derek Dubois	University of Rhode Island
Jin Fang	Clark University
Joy Field	Boston College
Renee Flasher	Penn State Harrisburg
Farnaz Ghashami	Drexel University
Douglas Hales	University of Rhode Island
Alisha Harper	Bellarmino University
Teray Johnson	Harrisburg University of Science and Technology
Haleh Karimi	Bellarmino University
Sue Kong	Kutztown University
EunSu Lee	New Jersey City University
Gang Li	Bentley University
Linda Boardman Liu	Boston College
Xiangrong Liu	Bridgewater State University
Dinesh Pai	Penn State Harrisburg
Fatma Pakdil	Eastern Connecticut State University
Raktim Pal	James Madison University
Vasileios Pavlopoulos	University of Alabama in Huntsville
Adem Pinar	Shenandoah University
Loreen Powell	Marywood University
Raza Rafique	Kean University
Ali Raoufi	University of Rhode Island
Pedro Reyes	Baylor University
Kamal Wada Ringim Sabo	Morgan State University
Piyal Sarkar	Penn State Harrisburg
Siamack Shojai	William Paterson University
Ju Myung (JM) Song	University of Massachusetts Lowell
Ping Wang	James Madison University

■ NEDSI SERVICE RECOGNITION TO PAST PROGRAM CHAIRS

Neil Desnoyers

**St. Joseph's University
Co-Program Chair, NEDSI 2024**

Xiangrong Liu

**Bridgewater State University
Co-Program Chair, NEDSI 2024**

Amit Arora

**University of the District of Columbia
Program Chair, NEDSI 2023**

Theologos Homer Bonitsis

**New Jersey Institute of Technology
Program Chair, NEDSI 2022**

Dinesh Pai

**Penn State Harrisburg
School of Business Administration
Program Chair, NEDSI 2021**

Effie Stavrulaki

**Bentley University
School of Business
Co-Program Chair, NEDSI 2020**

Gang Li

**Bentley University
School of Business
Co-Program Chair, NEDSI 2020**

Eric Stein

**Penn State Great Valley
School of Graduate Professional Studies
Program Chair, NEDSI 2019**

Douglas Hales

**University of Rhode Island
College of Business
Program Chair, NEDSI 2018**

■ NEDSI SERVICE RECOGNITION TO PAST PRESIDENTS

Theologos Homer Bonitsis

New Jersey Institute of Technology
Martin Tuchman School of Management
President of NEDSI, 2023-2024

Dinesh Pai

Penn State Harrisburg
School of Business
President of NEDSI, 2022-2023

Effie Stavroulaki

Bentley University
Co-President of NEDSI, 2021-2022

Gang Li

Bentley University
Co-President of NEDSI, 2021-2022

Eric W. Stein

Penn State Great Valley
School of Graduate Professional Studies
President of NEDSI, 2020-2021

Douglas Hales

University of Rhode Island
College of Business
President of NEDSI, 2019-2020

Neşet Hikmet

University of South Carolina
College of Engineering and Computing
President NEDSI, 2018



Saint Peter's
UNIVERSITY

THE WORLD NEEDS MORE PEACOCKS

Dive deep into cutting-edge technologies like AI, Natural Language Processing and Big Data Analytics with a

**Ph.D. in
Data Science**

Online and Hybrid
options available

Scan here to
learn more:



2641 John F. Kennedy Boulevard
Jersey City, NJ 07306



■ BEST PAPER AWARDS COMMITTEE

Joy Field (Chair)	Boston College
Amit Arora	University of the District of Columbia
Melike Hazal Can	Harvard University
Jonathan Daigle	Monmouth University
Neil Desnoyers	Saint Joseph's University
Douglas Hales	University of Rhode Island
Alisha Harper	Bellarmino University
Sue Kong	Kutztown University
Gang Li	Bentley University
Linda Boardman Liu	Boston College
Dinesh Pai	Penn State Harrisburg
Vasileios Pavlopoulos	University of Alabama in Huntsville
Fatma Pakdil	Eastern Connecticut State University
Raktim Pal	James Madison University
Adem Pinar	Shenandoah University
Loreen Powell	Marywood University
Pedro Reyes	Baylor University
Kamal Wada Ringim Sabo	Morgan State University
Piyal Sarkar	Penn State Harrisburg
Siamack Shojai	William Patterson University
Ju Myong (JM) Song	University of Massachusetts Lowell
Xiangrong Liu	Bridgewater State University

■ UNDERGRADUATE/MASTER STUDENT POSTER COMPETITION COMMITTEE

Jennifer A. Swanson (Chair)	Stonehill College
Kathleen Ferris-Costa	Bridgewater State University
Krista Hill	Babson College
Kellyann Berube Kowalski	UMass Dartmouth

■ NEDSI 2025 BEST PAPER AWARDS

Best Application of Theory Award

Using Natural Language Processing for Data-driven Risk Management
Haleh Valian, Biogen

Best Contribution to Theory Award

*A Comparative Look into Coping and Other Organizational Theories
to Explain Organizational Resilience*
Raktim Pal, James Madison University
Nezih Altay, DePaul University

Bryant University Best Paper Award in Supply Chain Management and Logistics

*Collaborative New Product Development: The Implications of Long-term vs. Short-term Contract
Choice on Cost-sharing in Vertical Channels*
Abhishek Roy, Temple University
Ganesh Balasubramanian, T A Pai Management Institute

Richard Briotta Best Paper Award in Knowledge Management/Strategy

From Pixels to Insights: Large Language Models for Automated Data Visualization Enhancement
Andres Fortino, New York University

David M. Levine Best Paper Award in Innovative Education

Ethics of AI in Education
Alexander Pelaez, Hofstra University
Amal Jacobsen, Progressive School of Long Island
Jim Samuel, Rutgers University
Kaushik Sengupta, Hofstra University

Best Ph.D. Student Paper Award

Quantum Machine Learning Approaches in Malware Analysis: An Extended Review
Philip Shen, Marymount University
Jennifer Ferreras-Perez, Marymount University

Best Overall Conference Paper Award

*Generating Sparse and Plausible Counterfactual Explanations with Spatial Diversity
and High Fidelity for Artificial Intelligence Decision Systems*
Nolan Talaei, University of Massachusetts Lowell
Asil Oztekin, University of Massachusetts Lowell
Hongwei Zhu, University of Massachusetts Lowell
Luvai Motiwalla, University of Massachusetts Lowell

■ UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

UNDERGRADUATE STUDENT RESEARCH POSTERS

A Practical Guide for Navigating the PTE Tax Landscape

Caitlin Sizemore, Alisha Harper
Bellarmine University, Louisville, KY, USA

Educational Growth and Student Access at the University of Southern Maine: Expanding the Modalities and Frequency of Course Offerings

Isabella Rossborough, Simon Gabaree
University of Southern Maine, Portland, ME, USA

Pixel Meets Algorithms; ARTificial Intelligence Revolutionzes Video Game Design

Devon Tolbert
York College of Pennsylvania, York, PA, USA

Digitizing Historical Logistics Industrial Maps and Studying the Growth and Development of Jersey City Ports Post World War

Ankith Anil Kadagadakai, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

High Scores, Dark Agendas: Radicalization Within Gaming Platforms

Nevin Ferry, Tamara Schwartz (mentor)
York College of Pennsylvania, York, PA, USA

Assessing the Security of Operational Technology Industrial Control Systems Controlled by Information Technology Assets

Conner Everett, Tamara Schwartz (mentor)
York College of Pennsylvania, York, PA, USA

Patient Privacy: An Ethical Analysis of Emerging Tech in Hospitals

Ian Burns, Tamara Schwartz (mentor)
York College of Pennsylvania, York, PA, USA

From Reps to Replicas: The Rise of Fitness Digital Twins

Zachary Deal, Tamara Schwartz (mentor)
York College of Pennsylvania, York, PA, USA

Investigating the Impact of Tariffs on Global Trade in the Trump 2.0 Era

Youngwoo Cho, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

Reducing Fresh Produce Waste Using Cellulosic Bio-Coating

Joselyn Reyes, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

Exploring the Town Management and University Partnership Impact on Community and Economic Development in Bridgewater, Massachusetts

David Alatorre
Bridgewater State University, Bridgewater, MA, USA

Deep Reinforcement Learning for the Heterogeneous Capacitated Vehicle Routing Problem

Yuanymeng Zhu¹, Dan Liu²

¹Department of Management, Kean University, Union, NJ, USA. ²Department of Management, Kean University, Union, NJ, USA

■ UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

Information Gain-Based MCDM Approach for Urban Critical Freight Corridor Identification using Spatial-Temporal Data Analysis

Yiyang Liu, Dan Liu
Kean University, Union, NJ, USA

A Dynamic Road Damage Prediction Framework Using XGBoost and ArcGIS: Enhancing Flood Disaster Management in New Jersey

Haoxiang Liu, Dan Liu
Department of Management, Kean University, Union, NJ, USA

Artificial Intelligence Policy in Lethal Autonomous Weapons Systems (LAWS) across the United States, China, and the European Union

Janel Baxter, Dr. Anshu Arora, Dr. Amit Arora, Dr. Pradeep Behera
University of the District of Columbia, Washington DC, USA

Social Robotics and Human-Robot Interaction in Healthcare

Nkechi Iweriebor, Dr. Anshu Arora, Dr. Amit Arora, Dr. Bushra Saeed
University of the District of Columbia, Washington DC, USA

Emotional Artificial Intelligence in Customer Service: Can AI Truly Replace Human Empathy?

Tamaya Edwards, Dr. Anshu Arora, Dr. Amit Arora
University of the District of Columbia, Washington DC, USA

Enhancing Agricultural Efficiency and Safety: The Synergy of AI, Robotics, and Human-Centered Smart Farming

Nicholas Brown, Dr. Amit Arora, Dr. Anshu Arora
University of the District of Columbia, Washington DC, USA

A Multi-Objective Optimization Framework for ESG-Constrained Portfolios using Advanced Diversification Techniques

Ignacio Pericet, Amit Arora, Anshu Arora
University of the District of Columbia, Washington DC, USA

The Determinants of AI Usage in the Workplace: A Comparative Study between Job-Market Entrants and Young Professionals

Leslie Lopez, Krupali Soni, Kayla Grant, Aaron van Klyton, Md. Golam Kibria
Kean University, Union, NJ, USA

Navigating the Blue Ocean and Economy: Renewable Energy Solutions and Sustainable Maritime Practices

Josh Esteban, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

Urban Planning: Implementing Accessibility in City Design

Eman Hermas, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

Efficacy of Phishing Remediation: An Undergraduate Honors Thesis

Chandra Dhimal, Michael Douglas
Millersville University, Millersville, PA, USA

Medical Malpractice: Legal, Insurance, and Ethical Implications. Inspired by a True Story.

Claire Coons
Millersville University, Millersville, PA, USA

Rich Picture Analysis for Health Policy

Akira Dasgupta¹, Kuan-Lung Chen²

¹Riverside High School, Leesburg, VA, USA. ²The George Washington University, Washington DC, USA

■ UNDERGRADUATE/GRADUATE RESEARCH POSTER PARTICIPANTS

Euthanasia Rates Among Companion Animals in Los Angeles County Compared to Income Level

Isabelle Darcy

Bridgewater State University, Bridgewater, MA, USA

Evaluating Student Experiences with the Class Waitlist System at NDSU

Samantha Hendrickson, Anna Kessel, Kaylyn Haug, Rylee Stotz, Alison Hinze

North Dakota State University, Fargo, ND, USA

GRADUATE STUDENT RESEARCH POSTERS

The Convergence of Generative AI and Quality Assurance: Enhancing Pharmaceutical Supply Chain Resilience.

Aoibhin Lavery, Elif Kongar

University of New Haven, West Haven, CT, USA

Resilient Supply Chains, Resilient Communities: An AI-Powered Emergency Response Framework

Aarya Kshetri, Nahian Karim, Sneha Anumandla, Elif Kongar

University of New Haven, West Haven, Connecticut, USA

Generative AI in Omni-Channel Marketing

Aarya Upadhyay, Ellen Belitzky

University of New Haven, West Haven, CT, USA

AI for Cost Reduction in B2C E-Commerce Supply Chains

Simran Harjani, Ellen Belitzky

University of New Haven, West Haven, CT, USA

Pitfalls of Over-Dependency on AI in Financial Markets

Nikhita Rane, Ellen Belitzky

University of New Haven, West Haven, CT, USA

Emotional Intelligence as a Measure of Potential Academic Success

Angelina Zuger, Kubra Banu, Joash Kawal, Kabhi Shanmugasundaram

Clark University, Worcester, MA, USA

Problems Faced by Indian Exporters to Export Psyllium Husks to the USA

Dharmi Shah, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Leadership and Team Sports

Denish Patel¹, EunSu Lee²

¹Dallas Baptist University, Dallas, TX, USA. ²New Jersey City University, Jersey City, NJ, USA

If they Can, Why Can't We? Comparing the United States' and European Union's Approaches to Artificial Intelligence Governance and Regulation

Muano Nevhudzholi, Mariam Sanogo, Dr. Anshu Arora, Anshu Arora

University of the District of Columbia, Washington DC, USA

Air Quality Improvements: A Catalyst for Better Respiratory Health

Sara Rehman, Anvitaa Patne

Clark University, Worcester, MA, USA

NJCU

EXPLORE GRADUATE PROGRAMS IN THE SCHOOL OF BUSINESS AT NEW JERSEY CITY UNIVERSITY

SCHOOL OF BUSINESS

School of Business faculty hail from around the world with a wealth of experience and expertise that translates to a business education grounded in real-world applications. Students gain discipline-related skills in accounting, finance, marketing, and management that make them desirable hires in an increasingly competitive business world. Just as critical, students develop a hunger for lifelong learning that will serve them throughout their life of their business and academic careers.

DEGREE PROGRAMS

MASTER OF BUSINESS ADMINISTRATION

Business Administration (MBA) with a specialization in:

Business Analytics (STEM Program)

Healthcare Management

Finance

Organizational Management & Leadership

Supply Chain & Maritime Port Management

MASTER OF SCIENCE

Accounting (M.S.)

Accounting Bridge Program (B.S./M.S.)

Business Analytics and Data Science (M.S.) (STEM Program)

Business Information Systems (M.S.) (STEM Program)

Financial Technology (M.S.) (STEM Program)



ACBSP
Accredited
Programs

Harvard Business School Online Collaboration

Graduate Business Essentials reflects a collaboration with Harvard Business School's digital learning initiative—Harvard Business School Online CORE. Students are taught by world-class Harvard Business School Online faculty with in-person classes led by NJCU instructors. These are best-in-class industry professionals who help to coach students through the online material as they are introduced to real-world scenarios. The course focuses on developing basic fluency in analytics, economics, and financial accounting.



Harvard
Business
School
Online

HOW TO APPLY

Get Started at:

njcu.edu/applygrad

Application Code Waiver

GOKNIGHTS

QUESTIONS?

Contact NJCU Admissions Today:

Phone/Voice: 201-200-3409

Text: 551-214-1181

Website: www.njcu.edu

Email: gradadmissions@njcu.edu

Office of University Admissions

2039 Kennedy Boulevard

Hepburn Hall, Room 206

Jersey City, N.J. 07305-1597

NJCU





Conference Program Schedule of Presentations

March 27-29, 2025

**Hershey Lodge
Hershey, PA**

Panel and Keynote Sessions

Panel: The Tariff Tightrope: Balancing Business and Supply Chain Stability

Thursday, March 27 - 8:00 am

Location: Wild Rose A, B

Chair: Dinesh Pai

Discussants:

- *Lou Paoletti, Senior Director, Supply Chain Services, Phoenix Contact
- *Olga Sass, Supervisor, International Trade Compliance & Transportation, Phoenix Contact
- *Richard Young, Distinguished Professor Emeritus of Supply Chain Management, Penn State Harrisburg
- *Nick Divan, Vice President of Financial Planning Analysis, D&H Distributing (via Zoom)

Join leaders from industry and academia as they discuss how tariffs imposed by the current administration are impacting business and supply chain management decision-making.

Scholarly Journals Editors' Panel

Friday, March 28 - 8:00 am

Location: Wild Rose A

Chair: Theologos Homer Bonitsis

Panelists:

- * Xenophon Koufteros (Co-Editor-in-Chief of Decision Sciences Journal)
- * Pedro Reyes (Co-Editor-in-Chief of Decision Sciences Journal of Innovative Education)
- * Arash Azadegan (Associate Editor of Journal of Supply Chain Management)
- * Sang-Heui Lee (Editor-in-Chief of Journal of Managerial Issues)

Meet the Editors of four scholarly journals. Each editor will make a brief presentation about their journal, followed by time for Q&A.

Panel: Transformative Role of AI in Decision-Making

Friday, March 28 - 10:00 am

Location: Wild Rose A, B

Chair: Dinesh Pai

Panelists:

- * Ed Lada, Jr. (President and CEO, Goodwill Keystone)
- * Abha Dawesar (Chief Digital Analytics Officer, Institutional Division, Vanguard)
- * Jim Prussack (General Manager—Infrastructure, Security, and Industry for Americas Global Partner Solutions, Microsoft)
- * Scot Case (Vice President, Sustainability and Corporate Responsibility, National Retail Federation)

Join industry leaders from Microsoft, Vanguard, the National Retail Federation, and Goodwill Keystone Area as they explore how data and analytics shape business strategy, decision-making, and culture. They will discuss the transformative role of AI in decision-making and how their industries are adapting to an evolving landscape.

NEDSI 2025 Keynote

Friday, March 28 - 5:00 pm

Location: Empire A, B, C

Keynote Speaker: Adriana Souza, Senior Finance Manager at D&H Distributing

Title: Gen AI and Supply Chain Resilience: Future-Proofing Through Innovation

Adriana will discuss the transformative effects of generative AI on creating resilient and flexible supply chains, while also sharing practical insights from her experience at the intersection of data science and distribution.

Thursday, March 27

7:30 am **Registration**
Confection Hall Lobby

8:00 am **PAN-1: Panel: The Tariff Tightrope: Balancing Business and Supply Chain Stability**
Wild Rose A, B
Chaired by: Dinesh Pai

Panel: The Tariff Tightrope: Balancing Business and Supply Chain Stability

Lou Paioletti¹, Olga Sass², Richard Young³, Nick Divan⁴
(¹Senior Director, Supply Chain Services, Phoenix Contact. ²Supervisor, International Trade Compliance & Transportation, Phoenix Contact. ³Distinguished Professor Emeritus of Supply Chain Management, School of Business Administration, Penn State Harrisburg. ⁴Vice President of Financial Planning Analysis, D&H Distributing)

9:00 am **Breakfast/Brunch (sponsored by Kean University)**
Empire A, B

10:00 am **OMR-1: Advances in OM/OR: Empirical Studies**
Magnolia A
Chaired by: Dan Liu

The Effect of Consumer Attributions and Relationships on Return Recovery Outcomes in the Apparel Industry

Dan Liu¹, Guangzhi Shang² (¹James Madison University, Harrisonburg, VA, USA. ²Florida State University,

Tallahassee, FL, USA)

Analyzing Bikeshare User Preferences Across Urban Zones: A Case Study of San Francisco's Zoning Influence on Ridership

Kaverinathan Dhanapal, Sanuj Kumar (New Mexico State University, Las Cruces, NM, USA)

AI Technologies in Supply Chain: Aligning Innovation with Market Demand

Mahsa Soltaninejad, Oliver Schaer, Seung-Lae Kim (Drexel University, Philadelphia, PA, USA)

10:00 am **SCM-1: Contemporary Topics in Supply Chain Management I**

Magnolia B

Chaired by: Eric Addeo

Regulating safer consumer products through supply chain mapping

Penina Orenstein¹, Elena Galkina² (¹Seton Hall University, South Orange, NJ, USA. ²California Department of Toxic Substances Control, CA, USA)

Multi-tier supplier selection using total cost of ownership and data envelopment analysis

Feras Saleh, Anjali Awasthi (Concordia University, Montreal, Quebec, Canada)

An Interdisciplinary Approach to The Synthesis of a Novel Circular Value Chain

Eric Addeo¹, Deborah Helman¹, David Walters² (¹DeVry University, Iselin, NJ, USA. ²Independent Researcher, Port Macquarie, Australia)

Thursday, March 27 (continued)

10:00 am **DSS-1: Technology-Supported Decision Making**

Magnolia C

Chaired by: Bryan Higgs

Strategic Interactions in Climate Action: A Game-Theoretic Analysis of City Managers and World Leaders

Sampson Addo, Bryan Higgs (University of the District of Columbia, Washington DC, USA)

Advanced Forecasting for RIA Firms: Machine Learning and Quantum Approaches

Damon Gray (Drexel University, Philadelphia, PA, USA)

Enhancing Heterogeneous Treatment Effect Estimation in Personalized Marketing: Integrating Feature Embeddings with Meta-Learners and Causal Forests

Rashedul Hasan, Asil Oztekin (University of Massachusetts Lowell, Lowell, MA, USA)

A Game-Theoretic Analysis of the Electric Vehicle Transition: Consumer, Business, and Grid Perspectives

Bryan Higgs (University of the District of Columbia, Washington, DC, USA)

10:00 am **IEC-1: Innovations in Classroom Instruction I**

Wild Rose A

Chaired by: Ronda Mariani

An Analysis of the Impact of Student Perceptions of an ERPsim Game on Their Performance

Sung Shim (Seton Hall University, South Orange, NJ, USA)

Protecting Genetic Privacy: Lessons from the 23andMe Data Breach

Amelia Olsson, Shaoping Zhao (Stockton University, Galloway, NJ, USA)

Redefining Academic Support: The Synergy Between GPT Assistants and Embedded Librarians in Online Classrooms

Ronda Mariani, Michael Coffta (Commonwealth University of Pennsylvania, Bloomsburg, PA, USA)

Alice in Borderland: The King of Diamonds Classroom Activity

Michael Douglas (Millersville University, Millersville, PA, USA)

11:40 am **Lunch (on your own)**

1:00 pm **AFE-1: Contemporary Issues in Accounting**

Magnolia B

Chaired by: Nathan Slavin

Understanding the Formation of Joint Ventures: ASU

2023-05 (Nathan Slavin¹, Jianing Fang² (¹Hofstra University, Hempstead, New York, USA. ²Kean University, Newark, NJ, USA)

Asset Impairment Created by Climate Change: New Challenges for Recording and Reporting

A. J. Stagliano (Saint Joseph's University, Philadelphia, PA, USA)

Increasing Student Engagement and Motivation in Tax Classes through Age-Based Scenarios and Practical Application (K Bryan Menk¹, Claire Stoner¹, Serafina Samosky¹, Alli Marchant²

¹Duquesne University, Pittsburgh, PA, USA. ²Grand

Canyon University, Phoenix, AZ, USA)

Cross-Border Accounting Scandals and Frauds

Ikechukwu Ndu (William Paterson University, Wayne, NJ, USA)

Similarities between the development of the Jewish calendar and generally accepted accounting principles (GAAP)

Dov Fischer (Brooklyn College, Brooklyn, NY, USA)

1:00 pm

CSI-1: Ethics, Healthcare, and Equity in Emerging Technologies

Wild Rose A

Chaired by: Douglas Hales

Digital Twinning Capital Costs (CAPx) and Operating Costs (OPx) within Process: An Emulation in a Manufacturing Company

Douglas Hales (The University of Rhode Island, Kingston, RI, USA)

Navigating AI Governance and Ethics: Strategies for the Future

Haleh Karimi, Alisha Harper (Bellarmine University, Louisville, KY, USA)

A Conceptual Framework for Integrating Emerging Technologies into Occupational Health and Safety Management Systems

Aida Haghighi¹, Afroz Moatari-Kazerouni², Gyeong Jun Lee¹ (¹Toronto Metropolitan University, Toronto, ON, Canada. ²Widener University, Philadelphia, PA, USA)

The Impact of Social Media and Technology on Empowering Saudi Women Entrepreneurs in National Social Projects: Analytical Research in Accordance to the Saudi 2030 Vision

Noha Alharbi (University of Bridgeport, Bridgeport, CT, USA)

Developing SMARTER Robots Engaged in Robotic Interventions for Students with Learning Disabilities

Anshu Arora, Amit Arora (University of the District of Columbia, Washington DC, USA)

Thursday, March 27 (continued)

1:00 pm

IEC-2: AI in Education

Magnolia C

Chaired by: Ram Misra

Learning from Experience: A ChapGPT Application in a Classroom

Anil Aggarwal, Cong Zhang, Danielle Fowler (University of Baltimore, Baltimore, MD, USA)

Bridging Academic Learning and Workforce Readiness: The Role and Value of Generative AI and Custom GPTs in Marketing Education.

Ronda Mariani¹, Loreen Powell² (¹Commonwealth University of Pennsylvania, Bloomsburg, PA, USA. ²Marywood University, Dunmore, PA, USA)

Exploring A Constructive Use of ChatGPT in the Classroom

Yang Li, Ram Misra (Montclair State University, Montclair, NJ, USA)

Smart Training for Smart Plants: VR and AI in Wastewater Treatment

Sofia M. Vidalis, Olanrewaju I. Ayodele (Penn State Harrisburg, Middletown, PA, USA)

Thursday, March 27 (continued)

1:00 pm **HSM-1: Healthcare and Data Analytics I**

Wild Rose B

Chaired by: Andres Fortino

Towards a Data Science Pattern Language

Tianyu Wang, [Andres Fortino](#), Roy Lowrance (New York University, New York, NY, USA)

Nurse Staffing in Rural Hospitals and Hospital Performance

[Dinesh R. Pai](#)¹, Esmail Bahalkeh² (¹Penn State Harrisburg, School of Business Administration, Middletown, PA, USA. ²University of New Hampshire, Health Management and Policy Department, Durham, NH, USA)

A Validation Analysis of the Resident Matching Algorithm

[Kartikeya Puranam](#), [Briance Mascarenhas](#) (Rutgers University, Camden, NJ, USA)

2:40 pm **Coffee Break (sponsored by James Madison University)**

Confection Hall Lobby

3:00 pm **SOB-1: Organizational Resilience and Change Management**

Magnolia A

Chaired by: Raktim Pal

From First Contact to Fatigue: The Hidden Toll of College Admissions Work.

[Richelyn Penn Mekile](#), [David Gefen](#) (Drexel University, Philadelphia, PA, USA)

The Evolution of Teleworking: Analyzing Critical

Success Factors for Sustained Excellence

[Kellyann Kowalski](#)¹, [Jennifer Swanson](#)² (¹University of Massachusetts Dartmouth, North Dartmouth, MA, USA. ²Stonehill College, South Easton, MA, USA)

The Three Pillars of Digital Transformation: Technology Integration, Platforms, and Skills

Ganesh Bhatt, [Samuel Ejiaku](#) (Morgan State University, Baltimore, MD, USA)

A Comparative Look into Coping and Other Organizational Theories to Explain Organizational Resilience

[Raktim Pal](#)¹, Nezh Altay² (¹James Madison University, Harrisonburg, VA, USA. ²DePaul University, Chicago, IL, USA)

3:00 pm **SCM-2: Emerging Topics in Supply Chain Management (invited session)**

Magnolia B

Chaired by: Piyal Sarkar

Collaborative New Product Development: The Implications of Long-term vs. Short-term Contract Choice on Cost-sharing in Vertical Channels

[Abhishek Roy](#)¹, Ganesh Balasubramanian² (¹Temple University, Fox School of Business, Philadelphia, PA, USA. ²T A Pai Management Institute, Manipal, KA, India)

Combating Greenwashing in Green Supply Chains: A Game-Theoretic Approach to Auditing Practices with Blockchain Technology

Shubham Gupta, [Vignesh Kesavaprakash](#) (New Jersey Institute of Technology, Newark, NJ, USA)

Impact of Cocreation under Supply Chain Encroachment Setting

Abhishek Roy¹, [Piya Sarkar](#)², Shubham Gupta³ (¹Temple

University, Philadelphia, PA, USA. ²Penn State University, Harrisburg, PA, USA. ³New Jersey Institute of Technology, Newark, NJ, USA

The Role of Generative AI in World-Class Logistics and Supply Chain Management

Angappa Gunasekaran (Penn State University, Middletown, PA, USA)

Motiwalla (University of Massachusetts Lowell, Lowell, MA, USA)

Dynamic Vehicle Routing with Social Media Integration in Disaster Relief

Emre Kirac (Christopher Newport University, Newport News, VA, USA)

Thursday, March 27 (continued)

3:00 pm **DSS-2: Artificial Intelligence I**

Magnolia C

Chaired by: Nolan Talaei

Development and Structure of an Introductory Graduate Course in AI for Business

AJ Raven (Drexel University, Philadelphia, PA, USA)

Tackling the Ambiguity Challenge with Generative Artificial Intelligence: AICMA, A Framework for Identification, Classification and Mitigation of Ambiguity

Kushal Patel¹, Jim Samuel¹, Alexander Pelaez², Shreyash Kalal¹ (¹Rutgers University, New Brunswick, NJ, USA. ²Hofstra University, Hempstead, NY, USA)

ESHRO: An Innovative Evaluation Framework for AI-Driven Mental Health Chatbots

Sahar Sherwani¹, Zarak Khan², Jim Samuel², Rajiv Kashyap³, Kushal Patel² (¹Rutgers University, New Brunswick, NJ, USA. ²Rutgers University, New Brunswick, NJ, USA. ³William Paterson University, Wayne, NJ, USA)

Generating Sparse and Plausible Counterfactual Explanations with Spatial Diversity and High Fidelity for Artificial Intelligence Decision Systems

Nolan Talaei, Asil Oztekin, Hongwei Zhu, Luvai

3:00 pm

CSI-2: Workshop: An Ocean 97 and Green Reef Foundation Aquaculture Digital Twin Adoption Experience for HBCU's in Aquaculture with a Net Zero Waste Activity Focus

Wild Rose A

Chaired by: Benjamin Branch

An Ocean 97 and Green Reef Foundation Aquaculture Digital Twin Adoption Experience for HBCU's in Aquaculture with a Net Zero Waste Activity Focus

Benjamin Branch¹, Jarvis Green^{2,3}, Chris Tate⁴, Patrick Menash⁵, Kimberly Milligan⁶, Gulnihal Ozbay⁶, Micheal Casson⁶, Neset Hikmet⁷ (¹Green Reef Foundation, Doha, Qatar, Qatar. ²Green Reef Foundation, Baton Rouge, LA, USA. ³Oceans 97, Baton Rouge, LA, USA. ⁴Red Hat, Boston, Massachusetts, USA. ⁵Southern University, Baton Rouge, LA, USA. ⁶Delaware State University, Dover, DE, USA. ⁷University of South Carolina, Columbia, SC, USA)

3:00 pm

MCB-1: AI, VR, and Digital Transformation in Marketing

Wild Rose B

Chaired by: Eklou Amendah

Comparing Virtual Reality and Website as Shopping Medium: Perspective from the S.O.R Model

Eklou R. Amendah¹, Archana Kumar² (¹University of Southern Maine, Manchester, NH, USA. ²Montclair State

University, Montclair, NJ, USA)

The Mediating Role of Digital Marketing Capabilities on Digital Marketing Performance in SMEs: An RBV Perspective

Khiem Dinh Mai¹, Ja-shen Chen² (¹Global Partners Service Center Corp., Taipei, Taiwan. ²College of Management, Yuan Ze University, Taoyuan, Taiwan)

6:00 pm **President's Reception (sponsored by University of Massachusetts Lowell, Manning School of Business)**
Empire A, B, C

Friday, March 28

7:30 am **Registration**
Confection Hall Lobby

7:30 am **Continental Breakfast (sponsored by Goodwill Keystone Area)**
Empire A, B

8:00 am **SOB-2: Analyzing Key Factors for Organizational Success**
Magnolia A
Chaired by: Parvinder Gupta

Filling in the Gaps: Missing Data Methods for Hofstede's LTO Cultural Dimension

Eugene Hahn¹, Jonathan Doh² (¹Salisbury University, Salisbury, MD, USA. ²Villanova University, Villanova, PA, USA)

The Effect of CEO and CFO Compensation Structures on the Quarterly Guidance Cessation

Decision

William Kline¹, Tom Adams², Richard Brown¹
(¹Pennsylvania State University Harrisburg, Middletown, PA, USA. ²La Salle University, Philadelphia, PA, USA)

Employee Involvement in Organizational Social Responsiveness Initiatives: Balance between Internal and External Empathy

Shanggeun Rhee, Glynnis Tan (Kean University, Union, NJ, USA)

8:00 am **SCM-3: Supply Chain Resilience**
Magnolia B
Chaired by: Afrooz Moatari-Kazerouni

Structuring Resilient Supply Chains: A New Blueprint for Network Design

Afrooz Moatari-Kazerouni, Yvonne Lederer Antonucci
(Widener University, Chester, PA, USA)

Indoor Farm Supply Chain Management

Eric W Stein^{1,2}, Afrooz Moatari-Kazerouni³ (¹Penn State, Malvern, PA, USA. ²Center of Excellence for Indoor Agriculture, Philadelphia, PA, USA. ³Widener University, Chester, PA, USA)

Collaborative Climates and Supply Chain Flexibility: Keys to Competitive Advantage

Yan Jin (Quinnipiac University, Hamden, CT, USA)

Resilience to Supply Chain Disruptions: An Investigation of the Roles of Innovativeness and Green Supply Chain Management Practices During the COVID-19 Pandemic

Lindsay Givens (Southern Arkansas University, Magnolia, AR, USA)

Northeast Decision Sciences Institute 2024 Annual Conference March 27-29, 2025 *All times in EDT*

Friday, March 28 (continued)

8:00 am **DSS-3: Artificial Intelligence II**
Magnolia C
Chaired by: Alexander Pelaez

Comparative Case Study on Implementing Generative AI in Medical Practices to Ease Documentative Overburden: A Sociotechnical Systems Perspective
Sri Ramesh Eevani, Rajiv Nag (Drexel University, Philadelphia, PA, USA)

Threat of AI - How Serious Can It Be?
Jinchang Wang (Stockton University, Galloway, NJ, USA)

Building Generative Artificial Intelligence (GAI) Utilization Expertise and Trust in GAI Output: A Parsimony Quantitative Model Approach to Predictors in the Software Development Industry Dynamics
Triumf Qosej, David Gefen (Drexel University, Philadelphia, PA, USA)

Ethics of AI in Education
Alexander Pelaez¹, Amal Jacobsen², Jim Samuel³, Kaushik Sengupta¹ (¹Hofstra University, Hempstead, NY, USA. ²Progressive School of Long Island, Merrick, NY, USA. ³Rutgers University, New Brunswick, NJ, USA)

8:00 am **PAN-2: Scholarly Journals Editors' Panel**
Wild Rose A
Chaired by: Theologos Homer Bonitsis

Scholarly Journals Editors' Panel
Xenophon Koufteros¹, Pedro Reyes², Arash Azadegan³, Sang-Heui Lee⁴ (¹Co-Editor-in-Chief of Decision Sciences Journal, College Station, TX,

USA. ²Co-Editor-in-Chief of Decision Sciences Journal of Innovative Education, Waco, TX, USA. ³Associate Editor of Journal of Supply Chain Management, Newark, NJ, USA. ⁴Editor-in-Chief of Journal of Managerial Issues, Pittsburg, KS, USA)

9:30 am **Undergraduate/Masters Student Research Poster Session**
Confection Hall Lobby
Chaired by: Jennifer Swanson

9:40 am **Coffee Break (sponsored by Penn State Great Valley, School of Business Administration)**
Confection Hall Lobby

10:00 am **AFE-2: Current Issues in International Finance**
Magnolia B
Chaired by: Theologos Homer Bonitsis

Exploring Corporate Capital Structure and Overleveraging in the Pharmaceutical Industry
Samar Issa¹, Hussein Issa² (¹Saint Peter's University, Jersey City, NJ, USA. ²Rutgers University, New Brunswick, NJ, USA)

A Leveraged Bootstrap Causality Approach to Investigate Relationships between the Consumer Inflation, Stock Market and Real Estate Market during March 2020 – Oct 2024
Howard Lee (Commonwealth University of Pennsylvania, Bloomsburg, PA, USA)

A Reexamination of Foreign Exchange Hedging by Multinational Enterprises
Augustine Arize¹, John Malindretos², Theodoros

Northeast Decision Sciences Institute 2024 Annual Conference March 27-29, 2025 All times in EDT

Stamatopoulos³, Chuanqian Zhang² (¹Texas A&M, Commerce, TX, USA. ²William Paterson University, Wayne, NJ, USA. ³University of West Attica, Athens, Greece)

NAFTA/UMSCA and Ford: An Analysis of Trade Agreement Impacts on an Automotive Giant

Haleigh Casey¹, Eric Kyper¹, Michael Douglas² (¹Lynchburg University, Lynchburg, VA, USA. ²Millersville University, Millersville, PA, USA)

Predictability of Financial Restatements: Conventional vs. Machine Learning

Khaled Abdou (Penn State University - Berks, Reading, PA, USA)

Friday, March 28 (continued)

10:00 am **PAN-3: Panel: Transformative Role of AI in Decision Making**

Wild Rose A, B

Chaired by: Dinesh Pai

Panel: Transformative Role of AI in Decision-Making

Ed Lada Jr.¹, Abha Dawesar², Jim Prussack³, Scot Case⁴

(¹President and CEO, Goodwill Keystone. ²Chief Digital Analytics Officer, Institutional Division, Vanguard.

³General Manager—Infrastructure, Security, and Industry for Americas Global Partner Solutions, Microsoft. ⁴Vice President, Sustainability and Corporate Responsibility, National Retail Federation)

11:40 am **Lunch (on your own)**

11:40 am **NEDSI Board Meeting (NEDSI board members only)**

Magnolia D

1:00 pm **SOB-3: Investigating Dynamics Influencing Organizational Performance**

Magnolia A

Chaired by: Henry Adobor

Managerial Dysfunctions and Exogenous Shocks: The Case of the Hotel Industry During the COVID-19 Pandemic

William Kline, Richard Brown (Pennsylvania State University Harrisburg, Middletown, PA, USA)

Learning and Performance: An Investigation

Parvinder Gupta (Indian Institute of Management Ahmedabad, Ahmedabad, Gujarat, India)

A Political Analysis Framework for Examining Power Dynamics in Open Strategy

Henry Adobor (Quinnipiac University, Hamden, CT, USA)

Unlock Success: How Organizational Resilience Bridges Resources and Performance

Huiling Liu (Arkansas State University, Jonesboro, AR, USA)

1:00 pm **AI/Machine Learning in Supply Chain Decision Making**

Magnolia B

Chaired by: Pedro Reyes

AI and Supply Chain Resilience: Benefits, Challenges, and Future

Pedro Reyes¹, Christopher Meyer¹, John Visich², Pat Jaska³ (¹Baylor University, Waco, TX, USA. ²Bryant University, Smithfield, RI, USA. ³University of Mary Hardin-Baylor, Belton, TX, USA)

Leveraging Machine Learning for Enhanced Supply Chain Decision-Making: A Study on Shipment Pricing and Mode Selection

Khushi Mehta (Hofstra University, Hempstead, NY, USA)

Generative AI in Supply Chain: A Systematic Review of Opportunities, Benefits, and Challenges

Musa Shaikh, Chandra Prakash, Pavan Mutha (University of the Cumberlands, Williamsburg, KY, USA)

Empowering Learning Business Analytics through ChatGPT

Rose Sebastianelli, Nabil Tamimi, Iordanis Petsas (University of Scranton, Scranton, PA, USA)

The Pitcher's Dilemma

Nicholas Patrichuk, Alexander Pelaez (Hofstra University, Uniondale, NY, USA)

Advancing Stock Price Prediction: A Hybrid LLM and Transformer-LSTM Framework with Multi-Task Learning

Rashedul Hasan, Asil Oztekin (University of Massachusetts Lowell, Lowell, MA, USA)

Friday, March 28 (continued)

1:00 pm **IEC-3: Workshop: Harnessing AI: Strategies for Enhancing Education and Research**

Magnolia C

Chaired by: Kathleen Ferris

Harnessing AI: Strategies for Enhancing Education and Research Success

Kathleen Ferris (Bridgewater State University, Bridgewater, MA, USA)

1:00 pm **BDA-1: Sustainability and Socioeconomic Applications in the Age of AI**

Wild Rose A

Chaired by: Iordanis Petsas

Paradoxes in Combining Inequalities

Jinchang Wang (Stockton University, Galloway, NJ, USA)

Factors that Affect American Savings in the 21st Century

Grace Gallagher (Hofstra University, Hempstead, NY, USA)

1:00 pm **CSI-3: Issues in Sustainability, Investments/Crypto, and Effects of Emerging Technologies**

Wild Rose B

Chaired by: Subhasish Dasgupta

Leveraging IoT for Environmental Sustainability: Enhancing Practices at Mercedes-Benz

Jessica Schwartz^{1,2,3,4,5}, Patricia Jones^{6,2}, Lori Farr^{1,5} (¹University of the Cumberlands, Williamsburg, KY, USA. ²Baker College, Owosso, MI, USA. ³APUS, Washington, DC, USA. ⁴Nexford University, Washington, DC, USA. ⁵Western Governors University, Salt Lake City, UT, USA. ⁶Wilmington University, Wilmington, DE, USA)

International Investments in Artificial Intelligence: Determinants and Effects

Kraiwinee Bunyaratavej (Delaware State University, Dover, DE, USA)

Prompt Engineer - The New Programmer

Bay Arinze, Samir Shah (Drexel University, Philadelphia, PA, USA)

**A Review of Information Systems and
Cryptocurrency Research**

Subhasish Dasgupta¹, Michael Chuang², Yuan Xue³
(¹George Washington University, Washington, DC,
USA. ²University of Illinois, Springfield, Springfield, IL,
USA. ³Elizabeth City State University, Elizabeth City,
NC, USA)

Friday, March 28 (continued)

2:00 pm **Undergraduate/Masters Student Research Poster Session**
Confection Hall Lobby
Chaired by: Jennifer Swanson

2:40 pm **Coffee Break (sponsored by University of the District of
Columbia, School of Business and Public Administration)**
Confection Hall Lobby

3:00 pm **SCM-5: Quantitative Modeling in Supply Chain
Management**
Magnolia A
Chaired by: Raza Rafique

**Humanitarian Logistics Network System Design using
Machine Learning Base on the Productivity-Driven
Clustering Evaluation Method**
Jae-Dong Hong (South Carolina State University,
Orangeburg, SC, USA)

**Data-Driven Supply Chains and Firm Performance:
Unveiling the Role of Supply Chain Integration in
Emerging Economies**
Raza Rafique (Kean University, Union, NJ, USA)

**Artificial Intelligence-Enhanced Stochastic
Optimization to Mitigate Supply Chain Disruptions**
Ge Guo (University of Baltimore, Baltimore, MD, USA)

Trade Credit Contracts under Weather Risk
Piyal Sarkar¹, Mohamed Wahab Mohamed
Ismail², Liping Fang² (¹Penn State University,
Harrisburg, PA, USA. ²Toronto Metropolitan University,
Toronto, Canada)

3:00 pm **AFE-3: Contemporary Issues in Finance**
Magnolia B
Chaired by: Ikechukwu Ndu

**Bitcoin and Lottery Stocks - Competition for The
Investor's Attention**
Tao Li¹, Wenxiu Nan² (¹SUNY New Paltz, New Paltz,
NY, USA. ²Salisbury University, Salisbury, MD, USA)

**The Spillover Effect of Dual Holdings on Analysts'
Information Environment**
Seung Won Lee¹, Heeick Choi², Sunghoon Joo³, Hakjoon
Song³ (¹Penn State Harrisburg, Middletown, PA,
USA. ²Sungkyunkwan University, Seoul, Korea,
Republic of. ³California State University, Dominguez
Hills, Carson, CA, USA)

**From Investigation to Incarceration: Beyond A Single
Agency Prosecuting Financial Institution Related
Crimes**
Renee Flasher, Olivia Bolinsky (Penn State Harrisburg,
Middletown, PA, USA)

**Economic Determinants of profitability of listed banks
in the United States: A Panel Data Approach**
Ikechukwu Ndu¹, Emmanuel Anoruo², Chiaku
Chukwuogor³ (¹William Paterson University, Wayne, NJ,
USA. ²Coppin State University, Baltimore, MD,

USA. ³Eastern Connecticut State University, Willimantic, CT, USA

Buy-Now-Pay-Later (BNPL) Lending: Innovation and Opportunity

Subramanian Narayanaswamy (New Jersey Institute of Technology, Newark, NJ, USA)

Friday, March 28 (continued)

3:00 pm **IEC-4: Innovations in Course Curriculum**

Magnolia C

Chaired by: Bea Chiang

Experiential Learning- Interdisciplinary Project Collaboration to Engage Students in Cost Accounting Class

Bea Chiang, Diane Bates, Kim Pearson, Monisha Pulimoon (The College of New Jersey, Ewing, NJ, USA)

Real Time Analytics System Design: An Integration of AWS Kinesis, Apache Pinot, and Apache Superset

Zhengzhong Shi (University of Massachusetts Dartmouth, Dartmouth, MA, USA)

What Should We be Teaching in UG Introduction to Information Systems?

Neil Desnoyers (Saint Joseph's University, Philadelphia, PA, USA)

Enhancing Decision Support Skills in Engineering Graduate Students through Game Theory

Bryan Higgs (University of the District of Columbia, Washington DC, USA)

What Do Students Say About HBCUs? A Preliminary Text-Mining Analysis

Cretson Dalmadge, Shih Chou (Winston-Salem State University, Winston-Salem, NC, USA)

3:00 pm

BDA-2: Decision-Making, Uncertainty, and Financial Insights of Analytics

Wild Rose A

Chaired by: Jin Fang

Advancing E-Waste Management and Fraud Prevention with Ensemble Machine Learning in Remanufacturing

Elif Kongar¹, Gazi Duman¹, Surendra Gupta² (¹University of New Haven, West Haven, CT, USA. ²Northeastern University, Boston, MA, USA)

A Machine Learning Algorithm Based on Max-Normalized Efficiency-Oriented Evaluation Method for Measuring Performance

Jae-Dong Hong (South Carolina State University, Orangeburg, SC, USA)

A Data Envelopment Analysis Framework for Urban Ranking and Criteria Prioritization

Jin Fang¹, Yasamin Salmani², Fariborz Partovi³ (¹Clark University, Worcester, MA, USA. ²Bryant University, Smithfield, RI, USA. ³Drexel University, Philadelphia, PA, USA)

Socioeconomic Determinants of Banking Term Deposit Subscriptions: A Multivariate Analysis Approach

Prakarsha Mitra (Hofstra University, Hempstead, NY, USA)

Examining Housing Instability and the Homelessness Issue in the United States: A Data-Driven Analysis

Shilpa Balan, Latika Kapoor, Kajal Bhandare (California State University, Los Angeles, Los Angeles, CA, USA)

3:00 pm **LES-1: Ethics in Action: Navigating Work, Technology, and Global Accountability**

Wild Rose B

Chaired by: Dawn Adams-Harmon

Creating Resilience for Minorities and Female Workers, and the Role of Flexibility in Work Environments: A Mixed Method Study

Dawn Adams-Harmon, Jung Ah (Claire) Yung, Aleisha Lewis (Kean University, Union, NJ, USA)

CrowdStrike Outage: The Social Responsibility of Technology

Nancy Sullivan¹, Ellen Belitzky² (¹Exact Changes, River Vale, NJ, USA. ²University of New Haven, West Haven, CT, USA)

Modern Slavery and Human Trafficking in Supply Chains

Wojtek Wolfe, Srinivasa Puranam (Rutgers University, Camden, NJ, USA)

5:00 pm **NEDSI 2025 Keynote**

Empire A, B, C

Presentation by: Adriana Souza, Senior Finance Manager at D&H Distributing

Title: Gen AI and Supply Chain Resilience: Future-Proofing Through Innovation

6:00 pm **Gala Dinner (sponsored by Penn State Harrisburg, School of Business Administration)**

Empire A, B, C

Saturday, March 29

7:15 am **NEDSI General Business Meeting**

Wild Rose B

7:30 am **Registration**

Confection Hall Lobby

7:30 am **Continental Breakfast (sponsored by New Jersey Institute of Technology)**

Empire A, B

8:00 am **SUS-1: New Trends in ESG and Sustainability in Supply Chains**

Magnolia A

Chaired by: Konstantia Litsiou

Plastic Recycling Fueled with Blockchain-Driven Tokenization: ESG Optimization and Implications

Aichih Chang¹, Nesreen El-Rayes², Jim Shi¹, Fuqin Zhou¹ (¹New Jersey Institute of Technology, Newark, NJ, USA. ²Georgia State University, Atlanta, GA, USA)

Exploring the Sustainable Development of Taiwan's Taoist Temple Culture through the System Dynamics Approach

Tung-Shan Liao, Shan-Hai Lin Lung (College of Management, Yuan Ze University, Taoyuan, Taiwan)

AI and RBV: Generating Resilient Supply Chains

Mark Lennon¹, Jerry Carbo², Ian Langella² (¹Pennwest University, California, PA, USA. ²Shippensburg University, Shippensburg, PA, USA)

Judgmental forecasting for decarbonization pathways in the European energy system

Ali Gamaleldin Karam¹, Mohamed Elshahhat Ali Khedr¹, Francesco Cariatì¹, Emanuele Giovanni Carlo

Ogliari¹, Dina (Konstantia) Litsiou², Vasileios Bougioukos³ (¹Politecnico di Milano, MILANO, Italy. ²Manchester Metropolitan University, Manchester, United Kingdom. ³Notre Dame London, London, United Kingdom)

Literature Review on Resource Efficiency and Value Chain Innovation in Semiconductor Industry

Xiangrong Liu¹, Samuel Serna-Otalvaro¹, Anu Agarwal², Pradnya Nagarkar², Lionel Kimerling², Ajay Gupta², Sajan Saini², Takahiro Ono², Beth Unger² (¹Bridgewater State University, Bridgewater, MA, USA. ²Massachusetts Institute of Technology, Cambridge, MA, USA)

Saturday, March 29 (continued)

8:00 am **HSM-2: Healthcare and Data Analytics II**

Magnolia B

Chaired by: Fatma Pakdil

A Meta Analysis of 30-Day Unplanned Readmission Estimation Models

Fatma Pakdil¹, Maling Ebrahimpour², Drew Zhang², Benjavan Upatising³, Yuehwen Yih⁴, Mahtab Kouhizadeh² (¹Eastern Connecticut State University, Willimantic, CT, USA. ²University of Rhode Island, Kingston, RI, USA. ³Center for Innovation and Transformational Change Department, UMass Memorial, Worcester, MA, USA. ⁴Purdue University, West Lafayette, IN, USA)

Customer and Employee Perceptual Congruence in Service Co-production: A Simulation Study

Ahmet Ozkul (University of New Haven, West Haven, CT, USA)

The Role of Social Networks in the Perceptions of Health Professionals about New Technology: The

Example of ESOGÜ Health and Practice Research Hospital

Neset Hikmet¹, Menderes Tarcan² (¹University of South Carolina, Columbia, South Carolina, USA. ²Eskisehir Osmangazi University, Eskisehir, Turkey)

Tech for Health: Understanding and Identifying Quality Health Information Online

Loreen Powell¹, Gwendolyn Powell², Chaza Abdul¹, Ronda Mariani³ (¹Marywood University, Scranton, PA, USA. ²Penn State University, State College, PA, USA. ³Commonwealth University, Bloomsburg, PA, USA)

8:00 am

IEC-5: Innovations in Classroom Instruction II

Magnolia C

Chaired by: Ross Malaga

Managing Product Innovation: Lessons from Aviation History

Ram Gopalan (Rutgers, The State University of New Jersey, Camden, NJ, USA)

Online Asynchronous Case Teaching in an AI World

Ellen Thomas (New Jersey Institute of Technology, Newark, NJ, USA)

Enhancing Operations Management Education with AI: Engaging Students with Innovative In-Class Activities

Jonathan Jackson (Providence College, Providence, RI, USA)

Innovations in the Undergraduate Spreadsheet Modeling Course

Ross Malaga, Rreze Zejnullahi (Montclair State University, Montclair, NJ, USA)

Saturday, March 29 (continued)

8:00 am **AFE-4: Topics in Accounting, Finance, and Economics**
Wild Rose A
Chaired by: Richard Muszynski III

**Enhanced Probability of Default Estimation with
Esscher Transform and EM Algorithm**
Sung Ik Kim (Louisiana State University Shreveport,
Shreveport, LA, USA)

**A New Data Envelopment Analysis Approach that
Ranks Mutual funds and Mutual Fund Companies**
Richard Muszynski III (Wilkes University, Wilkes-Barre,
PA, USA)

**Ethics of ChatGPT in CPE qualified assessment for
CPAs**
Dov Fischer (Brooklyn College, Brooklyn, NY, USA)

**Beyond Linearity: Modeling Key U.S. Economic
Indicators with Rational Curves**
Taoufik Meklachi, Brian Adams (Pennsylvania State
University Harrisburg, Middletown, PA, USA)

9:40 am **Coffee Break (sponsored by the University of Massachusetts
Dartmouth, School of Business Administration)**
Confection Hall Lobby

10:00 am **OMR-2: Advances in OM/OR: Modeling Studies**
Magnolia A
Chaired by: Janos Pinter

**Model Templates for Optimization Model
Development in Excel**

Janos Pinter (Rutgers University, Piscataway, NJ, USA)

**Optimal Charging Control for Autonomous Electric
Ride-Sharing Fleet**
Ehsan Mahyari (University of Scranton, Scranton, PA,
USA)

**A Clustering-based Approach for Items Storage and
Allocation Problem**
Reza Nafar, Satyaveer Singh Chauhan, Salim Lahmiri
(Concordia University, Montreal, Canada)

**Cyclic Routing Problem with Time Separation
Constraints**
Gang Li¹, Anantaram Balakrishnan² (¹Bentley University,
Waltham, MA, USA. ²The University of Texas at Austin,
Austin, TX, USA)

10:00 am **SCM-6: Contemporary Topics in Supply Chain
Management II**
Magnolia B
Chaired by: Shengbin Wang

Cross-Channel Effect in Omnichannel Retailing
Shengbin Wang (The College of New Jersey, Ewing, NJ,
USA)

**Enhancing Urban Logistics Resilience with Parcel
Locker-Assisted Dual-Mode Multi-Objective Vehicle
Routing**
Dan Liu (Kean University, Union, New Jersey, USA)

**Deterring Counterfeits through the Introduction of a
Store-Brand Product**
Jing Chen¹, Bintong Chen² (¹Dalhousie University,
Halifax, NS, Canada. ²University of Delaware, Newark,
DE, USA)

Supply Chain Contracts and Quality Investment with Blockchain Adoption

Min Wang (Drexel University, Philadelphia, PA, USA)

Saturday, March 29 (continued)

10:00 am **DSS-4: Natural Language Processing**

Magnolia C

Chaired by: Andres Fortino

DataSciLawBot: An LLM Chatbot for the Application of Data Science Industry Laws To Business Decisions

Sawyer Huang, Andres Fortino, Roy Lowrance (New York University, New York, NY, USA)

Using Natural Language Processing for Data-driven Risk Management

Haleh Valian (Biogen, Cambridge, MA, USA. Harvard University, Cambridge, MA, USA)

Design Science and Behavioral Science in Information Systems Research (From 2004 to 2024)

Toyin Asonibare, Ganesh Bhatt (Morgan State University, Baltimore, MD, USA)

From Pixels to Insights: Large Language Models for Automated Data Visualization Enhancement

Andres Fortino (New York University, New York, NY, USA)

Interdependence Between Stock Return, Social Media Sentiments, and Geopolitical Risks

Xiaorui Zhu¹, Michael Lash², Kyle Li³ (¹Towson University, Towson, MD, USA. ²University of Kansas, Kansas, KS, USA. ³Thomas Wootton High School, Rockville, MD, USA)

10:00 am

CSI-4: Cybersecurity, Privacy, Supply Chain, and Quantum Issues in Emerging Technologies

Wild Rose A

Chaired by: Jyothsna Harithsa

Quantum Machine Learning Approaches in Malware Analysis: An Extended Review

Philip Shen, Jennifer Ferreras-Perez (Marymount University, Arlington, VA, USA)

Blockchain Adoption in Supply Chain: A Trust Building Mechanism

Md Golam Kibria (Kean University, Union, NJ, USA)

Traffic Congestion as an Antecedent of Data Breaches.

Anjee Gorkhali¹, Asim Shrestha² (¹Susquehanna University, Selinsgrove, PA, USA. ²Denver University, Denver, CO, USA)

Navigating Uncertainty: Technology and Artificial Intelligence (AI) Sentiment in ESG Reports Pre- and Post-Pandemic

Jyothsna Harithsa¹, Kofi Arhin² (¹SUNY Geneseo, Geneseo, NY, USA. ²Lehigh University, Bethlehem, PA, USA)

10:00 am

MCB-2: Emerging Trends in Marketing Strategy, Branding, and Consumer Engagement

Wild Rose B

Chaired by: Estefany Galdamez

Customer Relationship Management for Intangible Goods: The Effects of Customer Commitment, Engagement, Responsiveness, and Reactance on Customer Rejection

Ameya Badve¹, Ramakrishnainh Bezawada¹, Karthik Sridhar² (¹Baruch College, New York, NY, USA. ²Baruch College, New York, NY, USA)

Consumer Perspective on Forced Labor: Product Price Tolerance

Estefany Galdamez (New Jersey Institute of Technology, Newark, NJ, USA)

Reactions to Competence- versus Warmth-based Brand Transgressions: Role of Political Orientation

Feisal Murshed¹, Tilottama Chowdhury² (¹Kutztown University of Pennsylvania, Kutztown, PA, USA. ²Quinnipiac University, Hamden, CA, USA)

Bridging the Gap in Marketing Operations: The Agile Campaign Managed Customer Framework.

Amirhossein Saghezchi, Seung-Lae Kim (Drexel University, Philadelphia, PA, USA)

Future of Decision Sciences: Complexity, Innovation, Sustainability

55th NEDSI Annual Meeting • Philadelphia, PA • April 9-11, 2026

Celebrating America's 250th Anniversary in Philadelphia!



MEETING ANNOUNCEMENT

The 2026 Northeast Decision Sciences Annual Meeting will take place April 9-11, 2026, in Philadelphia, PA. The three-day conference begins on Thursday and concludes at Saturday noon. Conference registrants are invited to attend the Welcome Reception, President's Reception, and Gala Dinner Award Ceremony.

CALL FOR PAPERS

Full papers, abstracts, and workshops are invited for, but not limited to, the topic areas listed. Undergraduate and Master level students may submit proposals for Poster Sessions. Sessions involving practitioners will be given consideration. Submissions will be blind refereed and accepted full papers will be published in the online Conference Proceedings, subject to author approval. By submitting a manuscript, the author certifies that it is not copyrighted or previously published, has not been presented or accepted for presentation at another professional meeting, and is not currently under review for presentation at another professional meeting.

At least one of the authors certifies that he/she intends to register for and attend the conference to present the paper if it is accepted.

All papers, abstracts, and posters must be submitted electronically by January 7, 2026, through the conference website. If you have proposals for panels, roundtable discussions, sessions, and workshops, please email them to the program chair. For additional information visit <https://nedsi.decisionsciences.org/>

OUTSTANDING PAPER AWARDS

Outstanding papers are eligible for the following awards:

- David M. Levine Best Paper Award in Innovative Teaching
- Richard Briotta Best Paper Award in Knowledge Management/Strategy
- Bryant University Best Paper Award in Supply Chain Management and Logistics
- Best Ph.D. Student Paper Award
- Best Contribution to Theory Award
- Best Application of Theory Award
- Best Overall Conference Paper Award

PROGRAM TRACKS

- Accounting, Finance, and Economics
- Big Data, Analytics, and Knowledge Management
- Cyber Security, IT, and Emerging Technologies
- DSS, Machine Learning, and AI
- Healthcare Analytics and Services Management
- Human-Technology Interface
- Innovative Education, Curriculum, and Cases
- Innovation and Creativity
- Legal, Ethical, and Social Issues
- Marketing and Consumer Behavior
- Operations Management/Operations Research
- Public Administration and Policy
- Strategy and Organizational Behavior
- Supply Chain Management and Logistics
- Sustainability Management
- Undergraduate/Masters Student Posters

HOTEL ARRANGEMENTS

The host hotel is the [Sonesta Philadelphia Rittenhouse Square](#). For reservations with the conference rate, a link will be made available on the conference website.

HOST SPONSOR



**Widener
University**

CONTACT INFORMATION

Afrooz Moatari-Kazerouni, Ph.D.
2026 NEDSI Program Chair
School of Business Administration
Widener University
One University Place, Chester, PA 19013
amoatarikazerouni@widener.edu

Thank you to our sponsors!

Grand Platinum level



PennState
Harrisburg

Platinum level



**MANNING SCHOOL
OF BUSINESS**

Gold level



KEAN





MARTIN TUCHMAN
SCHOOL OF MANAGEMENT

Silver level



Bronze level



Table of Contents

Full Papers

Accounting, Finance, and Economics

<u>Increasing Student Engagement and Motivation in Tax Classes through Age-Based Scenarios and Practical Application</u>	1

[K Bryan Menk](#)¹, [Claire Stoner](#)¹, Serafina Samosky¹, Alli Marchant²

¹Duquesne University, Pittsburgh, PA, USA. ²Grand Canyon University, Phoenix, AZ, USA

<u>NAFTA/UMSCA and Ford: An analysis of trade agreement impacts on an automotive giant</u>	14

Haleigh Casey¹, Eric Kyper¹, [Michael Douglas](#)²

¹Lynchburg University, Lynchburg, VA, USA. ²Millersville University, Millersville, PA, USA

Big Data and Business Analytics

<u>Paradoxes in Combining Inequalities</u>	23

[Jinchang Wang](#)

Stockton University, Galloway, New Jersey, USA

<u>Factors that Affect American Savings in the 21st Century</u>	29

[Grace Gallagher](#)

Hofstra University, Hempstead, NY, USA

<u>The Pitcher's Dilemma</u>	41

[Nicholas Patrichuk](#), [Alexander Pelaez](#)

Hofstra University, Uniondale, NY, USA

<u>A MACHINE LEARNING ALGORITHM BASED ON MAX-NORMALIZED EFFICIENCY-ORIENTED EVALUATION METHOD FOR MEASURING PERFORMANCE</u>	53

[Jae-Dong Hong](#)

South Carolina State University, Orangeburg, SC, USA

[Socioeconomic Determinants of Banking Term Deposit Subscriptions: A Multivariate Analysis Approach](#) 64

[Prakarsha Mitra](#)

Hofstra University, Hempstead, New York, USA

Cyber Security, IT, and Emerging Technologies

[Leveraging IoT for Environmental Sustainability: Enhancing Practices at Mercedes-Benz](#) 71

Jessica Schwartz^{1,2,3,4,5}, [Patricia Jones](#)^{6,2}, Lori Farr^{1,5}

¹University of the Cumberlands, Williamsburg, KY, USA. ²Baker College, Owosso, MI, USA. ³APUS, Washington, DC, USA. ⁴Nexford University, Washington, DC, USA. ⁵Western Governors University, Salt Lake City, UT, USA. ⁶Wilmington University, Wilmington, DE, USA

[Prompt Engineer - The New Programmer](#) 90

[Bay Arinze](#), [Samir Shah](#)

Drexel University, Philadelphia, PA, USA

[Quantum Machine Learning Approaches in Malware Analysis: An Extended Review](#) 100

[Philip Shen](#), [Jennifer Ferreras-Perez](#)

Marymount University, Arlington, VA, USA

[Navigating Uncertainty: Technology and Artificial Intelligence \(AI\) Sentiment in ESG Reports Pre- and Post-Pandemic](#) 110

[Jyothsna Harithsa](#)¹, [Kofi Arhin](#)²

¹SUNY Geneseo, Geneseo, New York, USA. ²Lehigh University, Bethlehem, PA, USA

Decision Support Systems (AI, ML, etc.) and Knowledge Management

[Threat of AI - How Serious Can It Be?](#) 124

[Jinchang Wang](#)

Stockton University, Galloway, New Jersey, USA

[Ethics of AI in Education](#) 131

[Alexander Pelaez](#)¹, [Amal Jacobsen](#)², Jim Samuel³, Kaushik Sengupta¹

¹Hofstra University, Hempstead, NY, USA. ²Progressive School of Long Island, Merrick, NY, USA. ³Rutgers University, New Brunswick, NJ, USA

<u>DataSciLawBot: An LLM Chatbot for the Application of Data Science Industry Laws To Business Decisions</u>	139
Sawyer Huang, <u>Andres Fortino</u> , Roy Lowrance NYU, NY, NY, USA	
<u>Design Science and Behavioral Science in Information Systems Research (From 2004 to 2024)</u>	163
<u>Toyin Asonibare</u> , Ganesh Bhatt Morgan State University, Baltimore, MD, USA	
<u>From Pixels to Insights: Large Language Models for Automated Data Visualization Enhancement</u>	193
<u>Andres Fortino</u> NYU, NYC, NY, USA	
<u>Towards a Data Science Pattern Language</u>	208
Tianyu Wang, <u>Andres Fortino</u> , Roy Lowrance NYU, NY, NY, USA	
<u>Development and Structure of an Introductory Graduate Course on AI for Business</u>	233
<u>AJ Raven</u> Drexel University, Philadelphia, PA, USA	
<u>Tackling the Ambiguity Challenge with Generative Artificial Intelligence: AICMA, A Framework for Identification, Classification and Mitigation of Ambiguity</u>	238
Kushal Patel ¹ , <u>Jim Samuel</u> ¹ , Alexander Pelaez ² , Shreyash Kalal ¹ ¹ Rutgers University, New Brunswick, New Jersey, USA. ² Hofstra University, Hempstead, New York, USA	
<u>ESHRO: An Innovative Evaluation Framework for AI-Driven Mental Health Chatbots</u>	252
Sahar Khan Sherwani ¹ , Zarak Khan ² , <u>Jim Samuel</u> ² , Rajiv Kashyap ³ , <u>Kushal Patel</u> ² ¹ Rutgers University, New Brunswick, New jersey, USA. ² Rutgers University, New Brunswick, New Jersey, USA. ³ William Paterson University, Wayne, New Jersey, USA	
<u>Advanced Forecasting for RIA Firms: Machine Learning and Quantum Approaches</u>	273
<u>Damon Gray</u> Drexel University, Philadelphia, PA, USA	

Healthcare and Services Management

[Tech For Health: Understanding and Identifying Quality Health Information Online](#) 296

[Loreen Powell¹](#), [Gwendolyn Powell²](#), Chaza Abdul¹, Ronda Mariani³

¹Marywood University, Scranton, PA, USA. ²Penn State University, State College, PA, USA. ³Commonwealth University, Bloomsburg, PA, USA

Innovative Education, Curriculum, and Cases

[Real Time Analytics System Design: An Integration of AWS Kinesis, Apache Pinot, and Apache Superset](#) 301

[Zhengzhong Shi](#)

University of Massachusetts Dartmouth, Dartmouth, MA, USA

[Managing Product Innovation: Lessons from Aviation History](#) 307

[Ram Gopalan](#)

Rutgers, The State University of New Jersey, Camden, NJ, USA

[Bridging Academic Learning and Workforce Readiness: The Role and Value of Generative AI and Custom GPTs in Marketing Education.](#) 334

[Ronda Mariani¹](#), [Loreen Powell²](#)

¹Commonwealth University of Pennsylvania, Bloomsburg, PA, USA. ²Marywood University, Dunmore, PA, USA

[Smart Training for Smart Plants: VR and AI in Wastewater Treatment](#) 338

Sofia M. Vidalis, [Olanrewaju I. Ayodele](#)

Penn State Harrisburg, Middletown, PA, USA

[Protecting Genetic Privacy: Lessons from the 23andMe Data Breach](#) 349

[Amelia Olsson](#), [Shaoping Zhao](#)

Stockton University, Galloway, New Jersey, USA

Marketing and Consumer Behavior

[The Mediating Role of Digital Marketing Capabilities on Digital Marketing Performance in SMEs: An RBV Perspective](#) 360

[Khiem Dinh Mai](#)¹, Ja-shen Chen²

¹Global Partners Service Center Corp., Taipei, Taiwan. ²College of Management, Yuan Ze University, Taoyuan, Taiwan

Strategy and Organizational Behavior

[The Three Pillars of Digital Transformation: Technology Integration, Platforms, and Skills](#) 402

Ganesh Bhatt, [Samuel Ejiaku](#)

Morgan State University, Baltimore, MD, USA

Supply Chain Management and Logistics

[Leveraging Machine Learning for Enhanced Supply Chain Decision-Making: A Study on Shipment Pricing and Mode Selection](#) 419

[Khushi Mehta](#)

Hofstra University, Hempstead, New York, USA

[HUMANITARIAN LOGISTICS NETWORK SYSTEM DESIGN USING MACHINE LEARNING BASED ON THE PRODUCTIVITY-DRIVEN CLUSTERING EVALUATION METHOD](#) 432

[Jae-Dong Hong](#)

South Carolina State University, Orangeburg, SC, USA

[Artificial Intelligence-Enhanced Stochastic Optimization to Mitigate Supply Chain Disruptions](#) 447

[Ge Guo](#)

University of Baltimore, Baltimore, MD, USA

Abstracts

Accounting, Finance, and Economics

[Bitcoin and Lottery Stocks - Competition for The Investor's Attention](#) 466

[Tao Li](#)¹, Wenxiu Nan²

¹SUNY New Paltz, New Paltz, NY, USA. ²Salisbury University, Salisbury, MD, USA

[The Spillover Effect of Dual Holdings on Analysts' Information Environment](#) 467

[Seung Won Lee](#)¹, Heeick Choi², Sunghoon Joo³, Hakjoon Song³

¹Penn State Harrisburg, Middletown, PA, USA. ²Sungkyunkwan University, Seoul, Korea, Republic of. ³California State University, Dominguez Hills, Carson, California, USA

[From Investigation to Incarceration: Beyond A Single Agency Prosecuting Financial Institution Related Crimes](#) 468

[Renee Flasher](#), Olivia Bolinsky

Penn State Harrisburg, Middletown, PA, USA

[Economic Determinants of profitability of listed banks in the United States: A Panel Data Approach](#) 469

[Ikechukwu Ndu](#)¹, Emmanuel Anoruo², Chiaku Chukwuogor³

¹William Paterson University, Wayne, New Jersey, USA. ²Coppin State University, Baltimore, Maryland, USA. ³Eastern Connecticut State University, Willimantic, Connecticut, USA

[Buy-Now-Pay-Later \(BNPL\) Lending: Innovation and Opportunity](#) 470

[Subramanian Narayanaswamy](#)

New Jersey Institute of Technology, Newark, NJ, US

[Exploring Corporate Capital Structure and Overleveraging in the Pharmaceutical Industry](#) 471

[Samar Issa](#)¹, Hussein Issa²

¹Saint Peter's University, Jersey City, NJ, USA. ²Rutgers University, New Brunswick, NJ, USA

[A Leveraged Bootstrap Causality approach to investigate relationships between the Consumer inflation, stock market and real estate market during March 2020 – Oct 2024 which cover the rise of the inflation and the decay.](#) 472

[Howard Lee](#)

Commonwealth University of PA, Bloomsburg, PA, USA

<u>A Reexamination of Foreign Exchange Hedging by Multinational Enterprises</u>	473
Augustine Arize ¹ , John malindretos ² , Theodoros Stamatopoulos ³ , Chuanqian Zhang ²	
¹ Texas A&M, Commerce, Texas, USA. ² William Paterson U, Wayne, NJ, USA. ³ University of West Attica, Athens, Greece	
<u>Predictability of Financial Restatements: Conventional vs. Machine Learning</u>	474
<u>Khaled Abdou</u>	
Penn State University - Berks, Reading, PA, USA	
<u>Enhanced Probability of Default Estimation with Esscher Transform and EM Algorithm</u>	475
<u>Sung Ik Kim</u>	
Louisiana State University Shreveport, Shreveport, LA, USA	
<u>A New Data Envelopment Analysis Approach that Ranks Mutual funds and Mutual Fund Companies</u>	476
<u>Richard Muszynski III</u>	
Wilkes University, Wilkes-Barre, PA, USA	
<u>Ethics of ChatGPT in CPE qualified assessment for CPAs</u>	477
<u>Dov Fischer</u>	
Brooklyn College, Brooklyn, NY, USA	
<u>Beyond Linearity: Modeling Key U.S. Economic Indicators with Rational Curves</u>	478
<u>Taoufik Meklachi</u> , Brian Adams	
Pennsylvania State University Harrisburg, Middletown, PA, USA	
<u>Understanding the Formation of Joint Ventures: ASU 2023-05</u>	479
<u>Nathan Slavin</u> ¹ , JIANING FANG ²	
¹ HOFSTRA UNIVERSITY, HEMPSTEAD, NEW YORK, USA. ² KEAN UNIVERSITY, NEWARK, NJ, USA	
<u>Asset Impairment Created by Climate Change: New Challenges for Recording and Reporting</u>	480
<u>A. J. Stagliano</u>	
Saint Joseph's University, Philadelphia, PA, USA	
<u>Cross-Border Accounting Scandals and Frauds</u>	481
<u>Ikechukwu Ndu</u>	
William Paterson University, Wayne, New Jersey, USA	

[Similarities between the development of the Jewish calendar and generally accepted accounting principles \(GAAP\)](#) 482

[Dov Fischer](#)

Brooklyn College, Brooklyn, NY, USA

Big Data and Business Analytics

[Empowering Learning Business Analytics through ChatGPT](#) 483

Rose Sebastianelli, Nabil Tamimi, [Iordanis Petsas](#)

University of Scranton, Scranton, Pa, USA

[Advancing Stock Price Prediction: A Hybrid LLM and Transformer-LSTM Framework with Multi-Task Learning](#) 484

[Rashedul Hasan](#), Asil Oztekin

University of Massachusetts Lowell, Lowell, MA, USA

[Advancing E-Waste Management and Fraud Prevention with Ensemble Machine Learning in Remanufacturing](#) 485

[Elif Kongar](#)¹, Gazi Duman¹, Surendra Gupta²

¹University of New Haven, West Haven, CT, USA. ²Northeastern University, Boston, MA, USA

[A Data Envelopment Analysis Framework for Urban Ranking and Criteria Prioritization](#) 486

[Jin Fang](#)¹, Yasamin Salmani², Fariborz Partovi³

¹Clark University, Worcester, MA, USA. ²Bryant University, Smithfield, RI, USA. ³Drexel University, Philadelphia, PA, USA

[Examining Housing Instability and the Homelessness Issue in the United States: A Data-Driven Analysis](#) 487

Shilpa Balan, [Latika Kapoor](#), [Kajal Bhandare](#)

California State University, Los Angeles, Los Angeles, CA, USA

Cyber Security, IT, and Emerging Technologies

[International Investments in Artificial Intelligence: Determinants and Effects](#) 488

[Kraiwinee Bunyaratavej](#)

Delaware State University, Dover, Delaware, USA

- [A Review of Information Systems and Cryptocurrency Research](#) 489
Subhasish Dasgupta¹, Michael Chuang², Yuan Xue³
¹George Washington University, Washington, DC, USA. ²University of Illinois, Springfield, Springfield, IL, USA. ³Elizabeth City State University, Elizabeth City, NC, USA
- [Blockchain Adoption in Supply Chain: A Trust Building Mechanism](#) 490
Md Golam Kibria
Kean University, Union, NJ, USA
- [Traffic Congestion as an antecedent of Data Breaches.](#) 491
Anjee Gorkhali¹, Asim Shrestha²
¹Susquehanna University, Selinsgrove, Pennsylvania, USA. ²Denver University, Denver, Colorado, USA
- [Navigating AI Governance and Ethics: Strategies for the Future](#) 492
Haleh Karimi, Alisha Harper
Bellarmine University, Louisville, KY, USA
- [A Conceptual Framework for Integrating Emerging Technologies into Occupational Health and Safety Management Systems](#) 493
Aida Haghighi¹, Afroz Moatari-Kazerouni², Gyeong Jun Lee¹
¹Toronto Metropolitan University, Toronto, ON, Canada. ²Widener University, Philadelphia, PA, USA
- [Developing SMARTER Robots Engaged in Robotic Interventions for Students with Learning Disabilities](#) 494
Anshu Arora, Amit Arora
University of the District of Columbia, Washington, DC, USA
- [An Ocean 97 and Green Reef Foundation Aquaculture Digital Twin Adoption Experience for HBCU's in Aquaculture with a Net Zero Waste Activity Focus](#) 495
Benjamin Branch¹, Jarvis Green^{2,3}, Chris Tate⁴, Patrick Menash⁵, Kimberly Milligan⁶, Gulnihal Ozbay⁶, Micheal Casson⁶, Neset Hikmet⁷
¹Green Reef Foundation, Doha, Qatar, Qatar. ²Green Reef Foundation, Baton Rouge, LA, USA. ³Oceans 97, Baton Rouge, LA, USA. ⁴Red Hat, Boston, Massachusetts, USA. ⁵Southern University, Baton Rouge, LA, USA. ⁶Delaware State University, Dover, DE, USA. ⁷University of South Carolina, Columbia, SC, USA

Decision Support Systems (AI, ML, etc.) and Knowledge Management

[Digital Twinning Capital Costs \(CAPx\) and Operating Costs \(OPx\) within Process: An Emulation in a Manufacturing Company](#) 496

[Douglas Hales](#)

The University of Rhode Island, Kingston, Rhode Island, USA

[Comparative Case Study on Implementing Generative AI in Medical Practices to Ease Documentative Overburden: A Sociotechnical Systems Perspective](#) 497

[Sri Ramesh Eevani, Rajiv Nag](#)

Drexel University, Philadelphia, PA, USA

[BUILDING GENERATIVE ARTIFICIAL INTELLIGENCE \(GAI\) UTILIZATION EXPERTISE AND TRUST IN GAI OUTPUT: A PARSIMONY QUANTITATIVE MODEL APPROACH TO PREDICTORS IN THE SOFTWARE DEVELOPMENT INDUSTRY DYNAMICS](#) 498

[Triumf Qosej](#), David Gefen

Drexel University, Philadelphia, PA, USA

[Using Natural Language Processing for Data-driven Risk Management](#) 499

[Haleh Valian](#)

Biogen, Cambridge, MA, USA. Harvard University, Cambridge, MA, USA

[Interdependence Between Stock Return, Social Media Sentiments, and Geopolitical Risks](#) 500

[Xiaorui Zhu](#)¹, Michael Lash², Kyle Li³

¹Towson University, Towson, MD, USA. ²University of Kansas, Kansas, KS, USA. ³Thomas Wootton High School, Rockville, MD, USA

[Generating Sparse and Plausible Counterfactual Explanations with Spatial Diversity and High Fidelity for Artificial Intelligence Decision Systems](#) 501

[Nolan Talaei, Asil Oztekin, Hongwei Zhu, Luvai Motiwalla](#)

University of Massachusetts Lowell, Lowell, MA, USA

[Dynamic Vehicle Routing with Social Media Integration in Disaster Relief](#) 502

[Emre Kirac](#)

Christopher Newport University, Newport News, VA, USA

[Strategic Interactions in Climate Action: A Game-Theoretic Analysis of City Managers and World Leaders](#) 503

Sampson Addo, [Bryan Higgs](#)

University of the District of Columbia, Washington DC, District of Columbia, USA

[Enhancing Heterogeneous Treatment Effect Estimation in Personalized Marketing: Integrating Feature Embeddings with Meta-Learners and Causal Forests](#) 504

[Rashedul Hasan](#), Asil Oztekin

University of Massachusetts Lowell, Lowell, MA, USA

[A Game-Theoretic Analysis of the Electric Vehicle Transition: Consumer, Business, and Grid Perspectives](#) 505

[Bryan Higgs](#)

University of the District of Columbia, Washington, DC, USA

Innovative Education, Curriculum, and Cases

[Harnessing AI: Strategies for Enhancing Education and Research Success](#) 506

[Kathleen Ferris](#)

Bridgewater State University, Bridgewater, MA, USA

[Experiential Learning- Interdisciplinary Project Collaboration to Engage Students in Cost Accounting Class](#) 507

[Bea Chiang](#), Diane Bates, Kim Pearson, Monisha Pulimoon

The College of New Jersey, Ewing, NJ, USA

[What should we be teaching in UG Introduction to Information Systems?](#) 508

[Neil Desnoyers](#)

Saint Joseph's University, Philadelphia, PA, USA

[Enhancing Decision Support Skills in Engineering Graduate Students through Game Theory](#) 509

[Bryan Higgs](#)

University of the District of Columbia, Washington, DC, USA

[What Do Students Say About HBCUs? A Preliminary Text-Mining Analysis](#) 510

[Cretson Dalmadge](#), Shih Chou

Winston-Salem State University, Winston-Salem, NC, USA

<u>Online Asynchronous Case Teaching in an AI World</u>	511
<u>Ellen Thomas</u>	
New Jersey Institute of Technology, Newark, NJ, USA	
<u>Enhancing Operations Management Education with AI: Engaging Students with Innovative In-Class Activities</u>	512
<u>Jonathan Jackson</u>	
Providence College, Providence, RI, USA	
<u>Innovations in the Undergraduate Spreadsheet Modeling Course</u>	513
<u>Ross Malaga, Rreze Zejnullahi</u>	
Montclair State University, Montclair, NJ, USA	
<u>Learning from Experience: A ChapGPT application in a classroom</u>	514
<u>Anil Aggarwal, Cong zhang, Danielle Fowler</u>	
University of Baltimore, Baltimore, MD, USA	
<u>Exploring A Constructive Use of ChatGPT in the Classroom</u>	515
<u>Yang Li, Ram Misra</u>	
Montclair State University, Montclair, NJ, USA	
<u>An Analysis of the Impact of Student Perceptions of an ERPsim Game on Their Performance</u>	516
<u>Sung Shim</u>	
Seton Hall University, South Orange, New Jersey, USA	
<u>Redefining Academic Support: The Synergy Between GPT Assistants and Embedded Librarians in Online Classrooms</u>	517
<u>Ronda Mariani, Michael Coffta</u>	
Commonwealth University of Pennsylvania, Bloomsburg, PA, USA	
<u>Alice in Borderland: The King of Diamonds Classroom Activity</u>	518
<u>Michael Douglas</u>	
Millersville University, Millersville, PA, USA	

Healthcare and Services Management

[A META ANALYSIS OF 30-DAY UNPLANNED READMISSION ESTIMATION MODELS](#) 519

Fatma Pakdil¹, Maling Ebrahimpour², Drew Zhang², Benjavan Upatising³, Yuehwern Yih⁴, Mahtab Kouhizadeh²

¹Eastern Connecticut State University, Willimantic, CT, USA. ²University of Rhode Island, Kingston, RI, USA. ³Center for Innovation and Transformational Change Department, UMass Memorial, Worcester, MA, USA. ⁴Purdue University, West Lafayette, IN, USA

[Customer and Employee Perceptual Congruence in Service Co-production: A Simulation Study](#) 520

Ahmet Ozkul

University of New Haven, West Haven, CT, USA

[The role of social networks in the perceptions of health professionals about new technology: The Example of ESOGÜ Health and Practice Research Hospital](#) 521

NESET HIKMET¹, MENDERES TARCAN²

¹University of South Carolina, Columbia, South Carolina, USA. ²Eskisehir Osmangazi University, Eskisehir, NA, Turkey

[Nurse Staffing in Rural Hospitals and Hospital Performance](#) 522

Dinesh R. Pai¹, Esmail Bahalkeh²

¹Penn State Harrisburg, School of Business Administration, Middletown, PA, USA. ²University of New Hampshire, Health Management and Policy Department, Durham, NH, USA

[A Validation Analysis of the Resident Matching Algorithm](#) 523

Kartikeya Puranam, Briance Mascarenhas

Rutgers University, Camden, NJ, USA

Legal, Ethical, and Social Issues

[Creating Resilience for Minorities and Female Workers, and the Role of Flexibility in Work Environments: A Mixed Method Study](#) 524

Dawn Adams-Harmon, Jung Ah (Claire) Yung, Aleisha Lewis

Kean University, Union, NJ, USA

[CrowdStrike Outage: The Social Responsibility of Technology](#) 525

[Nancy Sullivan](#)¹, [Ellen Belitzky](#)²

¹Exact Changes, River Vale, NJ, USA. ²University of New Haven, West Haven, CT, USA

[Modern Slavery and Human Trafficking in Supply Chains](#) 526

[Wojtek Wolfe](#), [Srinivasa Puranam](#)

Rutgers University, Camden, NJ, USA

Marketing and Consumer Behavior

[Customer Relationship Management for Intangible Goods: The Effects of Customer Commitment, Engagement, Responsiveness, and Reactance on Customer Rejection](#) 527

[Ameya Badve](#)¹, [Ramakrishnainh Bezawada](#)¹, [Karthik Sridhar](#)²

¹Baruch, NYC, NY, USA. ²Baruch, nyc, ny, USA

[Consumer Perspective on Forced Labor: Product Price Tolerance](#) 528

[Estefany Galdamez](#)

New Jersey Institute of Technology, Newark, NJ, USA

[Reactions to Competence- versus warmth-based brand transgressions: Role of political orientation](#) 529

[Feisal Murshed](#)¹, [Tilottama Chowdhury](#)²

¹Kutztown University of Pennsylvania, Kutztown, PA, USA. ²Quinnipiac University, Hamden, CA, USA

[Bridging the Gap in Marketing Operations: The Agile Campaign Managed Customer Framework.](#) 530

[Amirhossein Saghezchi](#), [Seung-Lae Kim](#)

Drexel University, Philadelphia, Pennsylvania, USA

[Comparing virtual reality and website as shopping medium: Perspective from the S.O.R model](#) 531

[Eklou R. Amendah](#)¹, [Archana Kumar](#)²

¹University of Southern Maine, Manchester, NH, USA. ²Montclair State University, Montclair, NJ, USA

Operations Management/Operations Research

- [Unlock Success: How Organizational Resilience Bridges Resources and Performance](#) 532
Huiling Liu
Arkansas State University, Jonesboro, AR, USA
- [Model Templates for Optimization Model Development in Excel](#) 533
Janos Pinter
Rutgers University, Piscataway, NJ, USA
- [Optimal Charging Control for Autonomous Electric Ride-Sharing Fleet](#) 534
Ehsan Mahyari
University of Scranton, Scranton, Pennsylvania, USA
- [A clustering-based approach for items storage and allocation problem](#) 535
Reza Nafar, Satyaveer Singh Chauhan, Salim Lahmiri
Concordia University, Montreal, Canada
- [Cyclic Routing Problem with Time Separation Constraints](#) 536
Gang Li¹, Anantaram Balakrishnan²
¹Bentley University, Waltham, MA, USA. ²The University of Texas at Austin, Austin, TX, USA
- [The Effect of Consumer Attributions and Relationships on Return Recovery Outcomes in the Apparel Industry](#) 537
Dan Liu¹, Guangzhi Shang²
¹James Madison University, Harrisonburg, Virginia, USA. ²Florida State University, Tallahassee, Florida, USA
- [Analyzing Bikeshare User Preferences Across Urban Zones: A Case Study of San Francisco's Zoning Influence on Ridership](#) 538
Kaverinathan Dhanapal, Sanuj Kumar
New Mexico State University, Las Cruces, New Mexico, USA
- [AI Technologies in Supply Chain: Aligning Innovation with Market Demand](#) 539
mahsa soltaninejad, Oliver Schaer, Seung-Lae Kim
drexel, Philadelphia, pen, USA

Strategy and Organizational Behavior

[Managerial Dysfunctions and Exogenous Shocks: The Case of the Hotel Industry During the COVID-19 Pandemic](#) 540

[William Kline](#), Richard Brown

Pennsylvania State University Harrisburg, Middletown, PA, USA

[Learning and Performance: An Investigation](#) 541

[Parvinder Gupta](#)

Indian Institute of Management Ahmedabad, Ahmedabad, Gujarat, India

[A Political Analysis Framework for Examining Power Dynamics in Open Strategy](#) 542

[Henry Adobor](#)

Quinnipiac University, Hamden, CT, USA

[Filling in the Gaps: Missing Data Methods for Hofstede's LTO Cultural Dimension](#) 543

[Eugene Hahn](#)¹, Jonathan Doh²

¹Salisbury University, Salisbury, MD, USA. ²Villanova University, Villanova, PA, USA

[The Effect of CEO and CFO Compensation Structures on the Quarterly Guidance Cessation Decision](#) 544

[William Kline](#)¹, Tom Adams², Richard Brown¹

¹Pennsylvania State University Harrisburg, Middletown, PA, USA. ²La Salle University, Philadelphia, PA, USA

[Employee involvement in organizational social responsiveness initiatives; balance between internal and external empathy](#) 545

[Shanggeun Rhee](#), [Glynnis Tan](#)

Kean University, Union, NJ, USA

[From First Contact to Fatigue: The Hidden Toll of College Admissions Work.](#) 546

[RICHELYN PENN MEKILE](#), [David Gefen](#)

Drexel University, Philadelphia, PA, USA

[The Evolution of Teleworking: Analyzing Critical Success Factors for Sustained Excellence](#) 547

[Kellyann Kowalski](#)¹, [Jennifer Swanson](#)²

¹University of Massachusetts Dartmouth, North Dartmouth, MA, USA. ²Stonehill College, South Easton, MA, USA

[A comparative look into coping and other organizational theories to explain organizational resilience](#) 548

[Raktim Pal](#)¹, Nezh Altay²

¹James Madison University, Harrisonburg, VA, USA. ²DePaul University, Chicago, IL, USA

Supply Chain Management and Logistics

[AI and supply chain resilience: benefits, challenges, and future](#) 549

[Pedro Reyes](#)¹, Christopher Meyer¹, John Visich², Pat Jaska³

¹Baylor University, Waco, Texas, USA. ²Bryant University, Smithfield, RI, USA. ³University of Mary Hardin-Baylor, Belton, Texas, USA

[Generative AI in Supply Chain: A Systematic Review of Opportunities, Benefits, and Challenges](#) 550

[Musa Shaikh](#), [Chandra Prakash](#), [Pavan Mutha](#)

University of the Cumberlands, Williamsburg, KY, USA

[Data-Driven Supply Chains and Firm Performance: Unveiling the Role of Supply Chain Integration in Emerging Economies](#) 551

[Raza Rafique](#)

Kean University, Union, NJ, USA

[Trade credit contracts under weather risk](#) 552

[Piyal Sarkar](#)¹, Mohamed Wahab Mohamed Ismail², Liping Fang²

¹Penn State University, Harrisburg, USA. ²Toronto Metropolitan University, Toronto, Canada

[Structuring Resilient Supply Chains: A New Blueprint for Network Design](#) 553

[Afrooz Moatari-Kazerouni](#), Yvonne Lederer Antonucci

Widener University, Chester, PA, USA

[Indoor Farm Supply Chain Management](#) 554

[Eric W Stein](#)^{1,2}, [Afrooz Moatari-Kazerouni](#)³

¹Penn State, Malvern, PA, USA. ²Center of Excellence for Indoor Agriculture, Philadelphia, PA, USA. ³Widener University, Chester, PA, USA

[Collaborative Climates and Supply Chain Flexibility: Keys to Competitive Advantage](#) 555

[Yan Jin](#)

Quinnipiac University, Hamden, CT, USA

- [Resilience to Supply Chain Disruptions: An Investigation of the Roles of Innovativeness and Green Supply Chain Management Practices During the COVID-19 Pandemic](#) 556
Lindsay Givens
Southern Arkansas University, Magnolia, Arkansas, US
- [Cross-channel effect in omnichannel retailing](#) 557
Shengbin Wang
The College of New Jersey, Ewing, NJ, USA
- [Enhancing Urban Logistics Resilience with Parcel Locker-Assisted Dual-Mode Multi-Objective Vehicle Routing](#) 558
Dan Liu
Kean University, Union, New Jersey, USA
- [Deterring counterfeits through the introduction of a store-brand product](#) 559
Jing Chen¹, Bintong Chen²
¹Dalhousie University, Halifax, NS, Canada. ²University of Delaware, Newark, DE, USA
- [Supply Chain Contracts and Quality Investment with Blockchain Adoption](#) 560
Min Wang
Drexel University, Philadelphia, PA, USA
- [Collaborative New Product Development: The Implications of Long-term vs. Short-term Contract Choice on Cost-sharing in Vertical Channels](#) 561
Abhishek Roy¹, Ganesh Balasubramanian²
¹Temple University, Fox School of Business, Philadelphia, PA, USA. ²T A Pai Management Institute, Manipal, KA, India
- [Combating Greenwashing in Green Supply Chains: A Game-Theoretic Approach to Auditing Practices with Blockchain Technology](#) 562
Shubham Gupta, Vignesh Kesavaprakash
New Jersey Institute of Technology, Newark, NJ, USA
- [Impact of cocreation under supply chain encroachment setting](#) 563
Abhishek Roy¹, Piyal Sarkar², Shubham Gupta³
¹Temple University, Philadelphia, USA. ²Penn State University, Harrisburg, USA. ³New Jersey Institute of Technology, Newark, USA

[The Role of Generative AI in World-Class Logistics and Supply Chain Management](#) 564

[Angappa Gunasekaran](#)

Penn State University, Middletown, PA, USA

[Regulating safer consumer products through supply chain mapping](#) 565

[Penina Orenstein](#)¹, Elena Galkina²

¹Seton Hall University, South Orange, NJ, USA. ²California Department of Toxic Substances Control, California, USA

[Multi-tier supplier selection using total cost of ownership and data envelopment analysis](#) 566

Feras Saleh, [Anjali Awasthi](#)

Concordia University, Montreal, Quebec, Canada

[An Interdisciplinary Approach to The Synthesis of a Novel Circular Value Chain](#) 567

[Eric Addeo](#)¹, [Deborah Helman](#)¹, David Walters²

¹DeVry University, Iselin, NJ, USA. ²Independent Researcher, Port Macquarie, Australia

Sustainability Management

[Plastic Recycling Fueled with Blockchain-Driven Tokenization: ESG Optimization and Implications](#) 568

[Aichih Chang](#)¹, Nesreen El-Rayes², Jim Shi¹, Fuqin Zhou¹

¹New Jersey Institute of Technology, Newark, NJ, USA. ²Georgia State University, Atlanta, GA, USA

[Exploring the Sustainable Development of Taiwan's Taoist Temple Culture through the System Dynamics Approach](#) 569

[Tung-Shan Liao](#), Shan-Hai Lin Lung

College of Management, Yuan Ze University, Taoyuan, Taiwan

[AI and RBV: Generating Resilient Supply Chains](#) 570

[Mark Lennon](#)¹, [Jerry Carbo](#)², [Ian Langella](#)²

¹Pennwest University, California, PA, USA. ²Shippensburg University, Shippensburg, PA, USA

[Judgmental forecasting for decarbonization pathways in the European energy system](#) 571

Ali Gamaleldin Karam¹, Mohamed Elshahhat Ali Khedr¹, Francesco Cariatì¹, Emanuele Giovanni Carlo Ogliari¹, [DINA \(KONSTANTIA\) LITSIOU](#)², Vasileios Bougioukos³

¹Politecnico di Milano, MILANO, Italy. ²MANCHESTER METROPOLITAN UNIVERSITY, MANCHESTER, United Kingdom. ³NOTRE DAME LONDON, LONDON, United Kingdom

[Literature Review on Resource Efficiency and Value Chain Innovation in Semiconductor Industry](#) 572

[Xiangrong Liu](#)¹, Samuel Serna-Otalvaro¹, Anu Agarwal², Pradnya Nagarkar², Lionel Kimerling², Ajay Gupta², Sajjan Saini², Takahiro Ono², Beth Unger²

¹Bridgewater State University, Bridgewater, MA, USA. ²Massachusetts Institute of Technology, Cambridge, MA, USA

Undergraduate Student Research Posters

[A Practical Guide for Navigating the PTE Tax Landscape](#) 573

[Caitlin Sizemore](#), Alisha Harper

Bellarmine University, Louisville, KY, USA

[Educational Growth and Student Access at the University of Southern Maine: Expanding the Modalities and Frequency of Course Offerings](#) 574

[Isabella Rossborough](#), [Simon Gabaree](#)

University of Southern Maine, Portland, Maine, USA

[Pixel Meets Algorithms; ARTificial Intelligence Revolutionizes Video Game Design](#) 575

[Devon Tolbert](#)

York College of Pennsylvania, York, PA, USA

[Digitizing historical logistics industrial maps and studying the growth and development of Jersey City Ports post world war](#) 576

[Ankith Anil Kadagadakai](#), EunSu Lee

New Jersey City University, Jersey City, NJ, USA

[High Scores, Dark Agendas: Radicalization Within Gaming Platforms](#) 577

[Nevin Ferry](#), Tamara Schwartz (mentor)

York College of Pennsylvania, York, PA, USA

- [Assessing the Security of Operational Technology Industrial Control Systems Controlled by Information Technology Assets](#) 578
Conner Everett, Tamara Schwartz (mentor)
 York College of Pennsylvania, York, PA, USA
- [Patient Privacy: An Ethical Analysis of Emerging Tech in Hospitals](#) 579
Ian Burns, Tamara Schwartz (mentor)
 York College of Pennsylvania, York, PA, USA
- [From Reps to Replicas: The Rise of Fitness Digital Twins](#) 580
Zachary Deal, Tamara Schwartz (mentor)
 York College of Pennsylvania, York, PA, USA
- [Investigating the Impact of Tariffs on Global Trade in the Trump 2.0 Era](#) 581
Youngwoo Cho, EunSu Lee
 New Jersey City University, Jersey City, NJ, USA
- [Reducing Fresh Produce Waste Using Cellulosic Bio-Coating](#) 582
Joselyn Reyes, EunSu Lee
 New Jersey City University, Jersey City, NJ, USA
- [Exploring the Town Management and University Partnership Impact on Community and Economic Development in Bridgewater, Massachusetts](#) 583
David Alatorre
 Bridgewater State University, Bridgewater, MA, USA
- [Deep Reinforcement Learning for the Heterogeneous Capacitated Vehicle Routing Problem](#) 584
Yuanyumeng Zhu¹, Dan Liu²
¹Department of Management, Kean University,, Union, New Jersey, USA. ²Department of Management, Kean University, Union, New Jersey, USA
- [Information Gain-Based MCDM Approach for Urban Critical Freight Corridor Identification using Spatial-Temporal Data Analysis](#) 585
YIYANG LIU, DAN LIU
 KEAN UNIVERSITY, UNION, NEW JERSY, USA
- [A Dynamic Road Damage Prediction Framework Using XGBoost and ArcGIS: Enhancing Flood Disaster Management in New Jersey](#) 586
Haoxiang Liu, Dan Liu
 Department of Management, Kean University, Union, New Jersey, USA

- [Artificial Intelligence Policy in Lethal Autonomous Weapons Systems \(LAWS\) across the United States, China, and the European Union](#) 587
Janel Baxter, Dr. Anshu Arora, Dr. Amit Arora, Dr. Pradeep Behera
University of the District of Columbia, Washington, District of Columbia, USA
- [SOCIAL ROBOTICS AND HUMAN-ROBOT INTERACTION IN HEALTHCARE](#) 588
Nkechi Iweriebor, Dr. Anshu Arora, Dr. Amit Arora, Dr. Bushra Saeed
University of the District of Columbia, Washington, District of Columbia, USA
- [Emotional Artificial Intelligence in Customer Service: Can AI Truly Replace Human Empathy?](#) 589
Tamaya Edwards, Dr. Anshu Arora, Dr. Amit Arora
University of the District of Columbia, Washington, District of Columbia, USA
- [Enhancing Agricultural Efficiency and Safety: The Synergy of AI, Robotics, and Human-Centered Smart Farming](#) 590
Nicholas Brown, Dr. Amit Arora, Dr. Anshu Arora
University of the District of Columbia, Washington, District of Columbia, USA
- [A Multi-Objective Optimization Framework for ESG-Constrained Portfolios using Advanced Diversification Techniques](#) 591
Ignacio Pericet, Amit Arora, Anshu Arora
University of the District of Columbia, Washington, DC, USA
- [The determinants of AI usage in the workplace: a comparative study between job-market entrants and young professionals](#) 592
Leslie Lopez, Krupali Soni, Kayla Grant, Aaron van Klyton, Md. Golam Kibria
Kean University, Union, NJ, USA
- [Navigating the Blue Ocean and Economy: Renewable Energy Solutions and Sustainable Maritime Practices](#) 593
Josh Esteban, EunSu Lee
New Jersey City University, Jersey City, NJ, USA
- [Urban Planning: Implementing Accessibility in City Design](#) 594
Eman Hermas, EunSu Lee
New Jersey City University, Jersey City, NJ, USA

[Efficacy of Phishing Remediation: An Undergraduate Honors Thesis](#) 595
[Chandra Dhimal](#), Michael Douglas
Millersville University, Millersville, PA, USA

[Medical Malpractice: Legal, Insurance, and Ethical Implications. Inspired by a True Story.](#) 596
[Claire Coons](#)
Millersville University, Millersville, PA, USA

[Rich Picture Analysis for Health Policy](#) 597
[Akira Dasgupta](#)¹, Kuan-Lung Chen²
¹Riverside High School, Leesburg, VA, USA. ²The George Washington University, Washington, DC, USA

[Euthanasia Rates Among Companion Animals in Los Angeles County Compared to Income Level](#) 598
[Isabelle Darcy](#)
Bridgewater State University, Bridgewater, MA, USA

[Evaluating Student Experiences with the Class Waitlist System at NDSU](#) 599
[Samantha Hendrickson](#), [Anna Kessel](#), [Kaylyn Haug](#), [Rylee Stotz](#), Alison Hinze
North Dakota State University, Fargo, ND, USA

Masters Student Research Posters

[The Convergence of Generative AI and Quality Assurance: Enhancing Pharmaceutical Supply Chain Resilience.](#) 600
[Aoibhin Lavery](#), Elif Kongar
University of New Haven, West Haven, CT, USA

[Resilient Supply Chains, Resilient Communities: An AI-Powered Emergency Response Framework](#) 601
[Aarya Kshetri](#), Nahian Karim, Sneha Anumandla, Elif Kongar
University of New Haven, West Haven, Connecticut, USA

[Generative AI in Omni-Channel Marketing](#) 602
[Aarya Upadhyay](#), Ellen Belitzky
University of New Haven, West Haven, CT, USA

<u>AI for Cost Reduction in B2C E-Commerce Supply Chains</u>	603
<u>Simran Harjani</u> , Ellen Belitzky University of New Haven, West Haven, CT, USA	
<u>Pitfalls of Over-Dependency on AI in Financial Markets</u>	604
<u>Nikhita Rane</u> , Ellen Belitzky University of New Haven, West Haven, CT, USA	
<u>Emotional Intelligence as a Measure of Potential Academic Success</u>	605
<u>Angelina Zuger</u> , <u>Kubra Banu</u> , <u>Joash Kawal</u> , <u>Kabhi Shanmugasundaram</u> Clark University, Worcester, MA, USA	
<u>PROBLEMS FACED BY INDIAN EXPORTERS TO EXPORT PSYLLIUM HUSKS TO THE USA</u>	606
<u>Dharmi Shah</u> , EunSu Lee New Jersey City University, Jersey City, NJ, US	
<u>LEADERSHIP AND TEAM SPORTS</u>	607
<u>Denish Patel</u> ¹ , EunSu Lee ² ¹ Dallas Baptist University, Dallas, TX, USA. ² New Jersey City University, Jersey City, NJ, USA	
<u>If they Can, Why Can't We? Comparing the United States' and European Union's Approaches to Artificial Intelligence Governance and Regulation</u>	608
<u>Muano Nevhudzholi</u> , Mariam Sanogo, Dr. Anshu Arora, Anshu Arora	
<u>Air Quality Improvements: A Catalyst for Better Respiratory Health</u>	609
<u>Sara Rehman</u> , <u>Anvitaa Patne</u> Clark University, Worcester, Massachusetts, USA	
<u>Author Index</u>	610

Full Papers



**Increasing Student Engagement and Motivation in Tax
Classes through Age-Based Scenarios and Practical Application**

K. Bryan Menk*
PhD, CPA (Va License #26374)
Associate Professor of Accounting
Palumbo-Donahue School of Business
Duquesne University
Rockwell Hall 815
600 Forbes Avenue
Pittsburgh, PA 15282
menkk@duq.edu

Claire Stoner
Master of Accountancy / BSBA Accounting
Palumbo-Donahue School of Business
Duquesne University

Serafina Samosky
Master of Accountancy / BSBA Accounting
Palumbo-Donahue School of Business
Duquesne University

Alli Marchant
Master of Accountancy
Grand Canyon University
BSBA Accounting
Palumbo-Donahue School of Business
Duquesne University

*Corresponding Author

ABSTRACT:

Teaching taxation often presents unique challenges to the instructor. While students have experienced paying taxes, most do not fully comprehend the wide range of laws underlying the tax system. To better help students learn and encourage student engagement, this paper proposes a redesigned method of instruction for teaching taxation. Instead of comprehensively teaching a particular tax concept (such as income or itemized deductions), the new method teaches topics based on the different stages of a person's life. This allows the class to begin with basic topics that increase in complexity and difficulty throughout the semester. Starting with rudimentary topics, allows students to develop confidence and skills that help them master the more challenging materials later in the course.

INTRODUCTION:

Taxation, a unique topic in academia, is one in which students already have some knowledge. Students are usually aware of taxes, from parents, news, advertisements, campaign fund drives, and even from required paperwork when beginning a new job. However, students often are not aware of details related to taxation. In order to better engage students and help them learn taxation we're encouraging a teaching approach that focuses on relating the material to students' lives.

This article describes a methodology of teaching an introductory tax course by using age-based scenarios, and related tax return preparations to promote student engagement in the learning process. The approach challenges students to learn by repeatedly applying conceptual material throughout the semester. They complete age-appropriate tax returns that include topics typically experienced at various stages in an individual's life. Students learn the standard tax topics in a manner in which they are integrated throughout the course instead of being compartmentalized, as they are in conventional methods. Students cement their understanding by applying this knowledge to complete a series of increasingly challenging tax returns. The initial assignments provide a basic introduction to tax topics while subsequent assignments weave into progressively complex tax situations. The class culminates with a final tax return that provides an all-inclusive review of the course topics. The emphasis on the continual application of concepts distinguishes this class from more traditional delivery approaches that employ few application-based assignments. The proposed process allows students to develop a deep conceptual understanding of taxation and application skills in a logical and engaging manner. Students are able to master basic analysis and preparation skills that will help in both their personal and professional lives.

Background on Experiential Learning:

Engaging and involving students can be a challenge in even the most interesting of subjects.

It is an area of focus and concern for educators across all topics of study and all levels of instruction (Carini, et al., 2006; Kahu, 2013; Kuh et al., 2008; Marks, 2000; Skinner and Belmont, 1993; Zhao and Kuh, 2004). The method discussed in this paper intends to increase a student's ability to understand and apply the information encountered in Individual Income Tax classes. Numerous studies have found that motivation is often a key to increasing student engagement and success (Ames, 1992; Greene et al., 2004; Komarraju and Karau, 2005; Linnenbrink and Pintrich, 2002; Scogin, 2016; Skinner and Belmont, 1993; Sun, 2014; Zumbrunn et al., 2014). As a result, instructors would benefit by adapting their classroom delivery to emphasize the relevance and applicability of course topics to students' lives. Additionally, the opportunity to collaborate, employ available resources, and become more involved with the application of topical information has been found to increase the level of student engagement and motivation (Dockter, 2010; Kahu, 2013; Greene et al., 2004; Marandos & Randall, 2012; Mega et al., 2014; Pintrich & Schrauben, 1992; Skinner & Belmont, 1993; Zumbrunn et al., 2014). While the idea of a "life cycle" for understanding taxation has already been considered and presented (Rubin, 1998; Kelliher, 2014), those approaches differs greatly from the one presented here.

Rubin (1998) presented a semester long case study involving a story line following an imaginary person. Rubin also incorporated elements that are broader in scope as they suggest economic decision making and planning instead of responding to a preset scenario. While, Rubin's process is similar, the differences are largely found in the application of concepts. In the method introduced in this manuscript, the students are analyzing the impact of the decisions already made, which is much more consistent with the work required in a compliance practice. The analysis and decision making is discussed in the program, but all data has been established before the students begin the returns. Rubin's method appears to be more analytical and decision focused instead of compliance and interpretation.

Kelliher (2014) created a graduate level program to track a married couple through the creation of investment opportunities and a business. His method focused more on the business impact on the taxpayers return and did not focus on this method as a comprehensive approach to introducing taxation through the appropriate ages. For example, in Kelliher's model, the taxpayers are already adults and married when the scenarios begin. In the current presentation, the tax payers are the students who are assuming a scenario as it applies to their personal lives at future ages. Dependent on the individualized nature of each assignment (age appropriate return), some are married, some are divorced, some have children, some are business owners. The idea of a life cycle may not be entirely novel, the application of the ideas of each age and the associated filing requirements rather the decision making process related to the financial transaction or the financial status in a business is completely new.

Motivation and engagement in courses are typically observed through increased participation in class discussions, active note taking in class, and development of a sense of ownership related to the completion of the required coursework. Such behaviors are indications of skills engagement, which are components of an effective learning strategy. As students become more actively engaged in their learning process, they develop

emotional engagement related to the materials. In addition, student performance on class assignments can also be seen as a factor influencing student engagement. When students perform well on assignments, they begin to feel that the topic is manageable and that they can succeed in the course (Floyd et al., 2009; Handelsman, 2005; Svanum and Bigatti, 2009, Webber et al., 2013).

Students' perceived value of a course also affects their engagement in it. When they can directly apply learned knowledge either to other courses or to future professional responsibilities, they can develop a personal connection, be more highly engaged in the course, and perform better (Floyd et al., 2009; Pintrich and de Groot, 1990; Zhao and Kuh, 2004; Zhao and Kuh, 2004). Linking assignments to students' real-world current and future experiences is one method of enhancing their understanding of the value of the course content. Encouraging experiential learning allows students to apply conceptual ideas and theories in a practical and personal manner, thus reinforcing the course's value to students. Interestingly, experiential learning has been found to be instrumental in increasing the engagement and success of students and has been defined and explained as Kolb's Cycle of Experiential Learning (Jelinek, 2017; Kolb and Kolb, 2009; Konak et al., 2014; Petkus, 2000).

The Cycle includes the following steps Abstract Conceptualization, Active Experimentation, Concrete Experience, and Reflective Observation. Abstract Conceptualization creates the basic understanding of topics by modifying perceptions of information into abstract concepts. Active Experimentation includes testing ideas and refining skills through experience. Concrete Experience allows students to engage directly in a realistic application of practical skills. Reflective Observation allows students to relate prior experiences to the modification of conceptual knowledge (Kolb, 1984). The ability of students to apply concepts through an experience-based process or project-based learning has been shown to increase their understanding of the topics (Bell, 2010; Blumenfeld et al., 1991; Kraiger et al., 1993; Roehl, 2013). This process allows students to learn concepts, experiment with the knowledge, apply concepts in a realistic setting, and modify their understanding through the application of the tools.

COMPARISON OF TAX CURRICULUM:

Taxation is a required course for all students who seek to earn an accounting degree and the topic comprises a significant portion of the Regulation section of the CPA exam (Uniform CPA Exam, 2017). While there has recently been the adjustment to the CPA exam to allow for a more tracked or directed focus called the Discipline Exams. Even with these discipline based focuses, there is still the requirement to tax the REG Section as part of the core. Despite the changes and more focused attention of the CPA exam, there is still and will always be the need to be competent and aware of the tax implications each business transaction and plan. The age-based scenarios personalize the learning experience for students and we believe it can contribute to enhanced levels of student engagement. The method focuses on how taxation relates to the student's daily lives. In addition, crafting classroom discussions around different taxes and their impact on students' decisions engages students in the learning process.

The American Institute of Certified Public Accountants (AICPA) has released a Model Tax Curriculum which provides information on the importance of taxation in an undergraduate curriculum. The model created a series of learning objectives that focus on tax research, policy considerations, statutory/regulatory/ ethical obligations, basic tax knowledge for individuals and businesses, the impact of taxation on economic decisions, etc. (AICPA, 2014). While these topics are relevant and important to the development of a complete understanding of taxation, the model curriculum does not include guidance on how to effectively teach the topics to students. Nor does the curriculum expand sufficiently to cover all the topics that would be encountered in the Discipline Exam (Tax Compliance and Planning). The curriculum carries too much material to be effectively covered in one course, but it does emphasize the need to have a complete and thorough understanding of the individual income tax rules and regulations. This would necessarily include a comprehensive ability to apply the conceptual ideas to the completion of or at least a comprehension of the tax return impact of the decisions being made in the course of business.

Based on personal experiences and a review of many syllabi of tax courses from a nationwide sample of universities and colleges, existing methods of teaching taxation normally include the use of a single (or sometimes two) tax return(s) that are comprehensive and include all course information in a real-world application example.

Additionally, numerous schools whose syllabi were examined focus on tax research but do not include practical application assignments. While the ability to perform tax research may be important, it does little to develop practical knowledge. In terms of Kolb's Cycle of Experiential Learning, this approach exemplifies only Abstract Conceptualization as students are required to research tax questions for which they lack any practical understanding. Some schools approached Kolb's Active Experimentation and Concrete Experience steps with a very small scope of practical engagement by including tax returns, but as a very limited portion of the coursework. Most schools had only one very complex return due at the end of the semester.

While some steps of Kolb's Cycle of Experiential Learning are employed in these courses, the cycle is incomplete and there are untapped opportunities for improving the student comprehension and learning. This paper proposes and describes an alternative instructional method, one that emphasizes a life cycle approach. This teaching methodology focuses on the preparation of tax returns related to a particular time in a person's life and utilizes research topics and related tax returns that increase in difficulty throughout the course. Using individual scenarios and data sets, students apply the concepts learned in class to real-life situations (Rubin, 1998).

After reviewing the syllabi, the authors examined textbooks from numerous major publishers (Pearson, McGraw-Hill, Wiley, etc.). The structuring of topics tends to focus on major groups (income in all aspects, and then deductions, etc.), with all related sub-topics

presented at once. As a result, students are unable to complete a return until much later in the semester.

As previously discussed, research on teaching engagement indicates that the use of relatable materials, numerous attainable assignments (increasing in difficulty throughout the semester), and emphasis on the value of the topic are significantly related to student's engagement in the class. However, the review of textbooks and syllabi from numerous schools do not indicate that these considerations have been a significant component of the educational design. The syllabi largely tend to follow the textbook design rather than applying the guidance found through engagement research.

METHODOLOGY:

In an effort to design a model for teaching and learning taxation that includes methods found to be supportive of higher student engagement and related learning, the concept of age-based scenarios as a teaching tool was created

The proposed age-based methodology was designed to engage students by first teaching tax concepts and then reinforcing them with immediate application. Confidence is built as understanding takes place, and new steps are more easily learned. The initial return and related research memo require basic knowledge. Each subsequent return increases in difficulty and complexity, however; the students will have gained confidence and can better complete later assigned complicated returns.

This teaching method intends to increase engagement, motivation, and positive student performance through extensive application of the concepts related to events that occur at selected ages throughout life. These ideas are introduced as events in the lives of the students. The topics cover the life statuses of individuals at ages 22, 27, 40, 55, and 68. Additionally, students are required to complete a short research memo (structured as a client letter or client file memo) on the topics relevant to each tax return.

As the students progress in the course, the tax situations become more complex. The "life" assigned tax returns reflect events appropriate for each given age. The instructor needs to manage the level of difficulty of successive assignments so that students are appropriately challenged while building their confidence in the process. By increasing the difficulty level as the semester progresses and the students' skillsets increase, students can discover better methods to apply the concepts to assignments in an effective manner (Brophy, 2010; Garis et al., 2002; Marandos and Randall, 2012; Sweller, 1994).

EXAMPLES OF AGE-APPROPRIATE TOPICS

Age 22:

Return Preparation: Filing status, tax rates and calculation, standard deduction, wages, interest earned, and federal and state withholdings.

Age 27:

Return Preparation: Marriage, dependency, salary advancement, municipal interest, capital gains and losses, dividends (ordinary and qualified), and student loans.

Age 40:

Return Preparation: Itemized deductions, home ownership (including mortgages and real estate taxes), child tax credit, childcare tax credit, divorce/alimony, child support, partnership income, charitable contributions, and casualty loss.

Age 55:

Return Preparation: Elderly dependent, own their own business, have partnership income, children are attending college (American Opportunity Credit, Lifetime Learning Credit), fringe benefits are emphasized (FSA, HSA, life insurance, etc.), and investment issues (AMT).

Age 68:

Return Preparation: Social Security benefits, gift taxes and estate planning, passive loss limitation rules, business loss limitation, annuity taxation, penalties on underpayment of estimated taxes, and interest on underpayments.

In the age 22 tax return, deals with wages, interest earned, federal and state withholdings, and scholarships. In addition, the standard deduction is introduced, and students are required to calculate their tax liability. These topics are sufficient to complete a simple tax return by using only the most basic of tax forms (Form 1040 – Individual Income Tax Return and the Schedule B – Interest and Ordinary Dividends). Not only does this initial assignment build students' confidence in their ability to address more ambiguous tax issues, but it also fosters a genuine interest in learning at an early stage. Importantly, this experience provides students with a tax foundation that better enables them to comprehend the increasingly complex tax situations that they will face in the ensuing years.

This initial return is basic enough to be assigned very early in the semester (usually during the third-course meeting of a two-day-a-week class schedule). The instructor needs to review the process of completing the required forms before assigning the next return.

In addition to the tax returns, numerous research assignments are assigned throughout the semester. Most of these relate specifically to the topics covered in a tax return.

Examples of Research Topics

Age 22:

Research: Taxable interest, earned wages, and standard deduction.

Age 27:

Research: Capital gain treatment, and dependents.

Age 40:

Research: Itemized deductions, childcare and child tax credits, interest paid deductions, partnership taxation, divorce topics, sole proprietorships, charitable contributions, casualty losses, and self-employment taxes.

Age 55:

Research: Alternative Minimum Tax (AMT), fringe benefits, and higher education credits.

Age 68:

Research: Gift taxes, business loss limitation, and hobby vs business losses.

Additionally, the instructor who developed the course uses a system that creates unique individual tax returns for each student (Menk and Malone, 2015). With the use of individualized tax return data, students will have a different number of children, wage amount, etc.

Overall, the returns prepared by students must be more complex and difficult with each passing age as new topics are regularly introduced, and each new return requires students to use skills developed and knowledge acquired in previous returns. This teaching methodology encourages students to focus on the process of learning by initially introducing new ideas, experimenting with the comprehension of those concepts, actively applying the concepts to practice, and reflecting on the experience. This approach was designed to utilize Kolb's Cycle of Experiential Learning throughout the learning process (Kolb, 1984).

Kolb's cycle, usually presented as a circular process in which the student is continuously learning and developing new skills, is assumed to have a starting point for the purposes of this discussion. To exhibit the use of Kolb's Cycle in the method, we have presented the Concrete Experimentation as the starting point and the cause of differentiation from other teaching methods. Students apply the rules and regulations of taxation to real-life scenarios using real tax forms. This new experience allows the students to understand the process and to develop skills and confidence. The students then apply the Reflective Observation step through the writing assignment in which they discuss their understanding of a topic relevant to the return. This assignment requires students to research the topic as it helps them develop a deeper understanding of the elements included in the tax return. The third step, Abstract Conceptualization, involves the introduction of new concepts and revisiting previously discussed ideas to encourage a more complex analysis of the topic. This level allows the students to develop a general understanding of the new topics and to create a broader understanding of the taxation system, both application and conceptual theory. The final step, Active Experimentation, occurs when the students review the previous tax return (graded and returned before assigning the next return) and begin to plan for the next more complex return. They

incorporate skills previously learned while using new forms, new elements of taxable income, and new concepts (Kolb, 1984).

SUMMARY:

This method is different from traditional teaching methods because it consciously integrates the steps of Kolb's Cycle of Experiential Learning. By using scenarios and tax forms to put conceptual ideas into practice numerous times in a semester, the students are able to employ each stage of the cycle multiple times in a single semester. The use of practical application early in the semester is a particularly significant departure from other methods of teaching. It promotes a more comprehensive understanding of concepts through regular application, thoughtful reflection, and the opportunity to refine the skills and understanding of the concepts.

The challenge with any change in method is to ensure that the change is warranted and produces the desired results. The age-based life cycle methodology described in the paper has been applied in numerous sections of Individual Income Tax over a multi-year span. Instructors' using this method have had numerous informal conversations with students, while enrolled in the class and after the completion of the course, to discuss the impact of the teaching method. The students' responses have been universally very supportive of the age-based method and the students have stated that they felt well prepared for the Regulation section of the CPA exam. The Regulation Section tests all topics in the tax returns age scenarios above and has task-based simulations that require test takers to fill in blank values on certain tax forms based on the information given.

Additionally, as a result of methodological change, the students were observed to be significantly more engaged in the class, and a drastic increase in the number of students who took advantage of the instructors' office hours.

This increase in commitment has created an opportunity to challenge the students through an increase in the workload and difficulty of the assignments. Regularly reinforcing the critical skill of completing a tax return through practical application has enhanced students conceptual understanding of the course material. Instead of the two or three very complex returns in the traditional method, the students complete five returns throughout the semester under the new method. The ease of the first return has created a situation in which the students develop a much higher level of confidence in their skills. This increase in confidence and student learning is reflected in their ability to complete the most complex returns and a positive attitude towards these assignments.

Furthermore, this approach could be suitable in an area such as management. Starting from a small start-up business to a complex international business with thousands of employees.

Regardless of the topic being taught, the concept of applying the materials to the students' lives at different stages should be considered as a means of increasing engagement and student learning. Unlike more traditional approaches to teaching taxation, course concepts are regularly reinforced as students apply their knowledge to numerous, increasingly complex tax returns throughout the semester.

REFERENCES

- AICPA (2014). Model Tax Curriculum. Aicpa.org.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, 84(3), 261.
- Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing house*, 83(2), 39-43.
- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational psychologist*, 26(3-4), 369-398.
- Brophy, J. E. (2013). Motivating students to learn. Routledge.
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in higher education*, 47(1), 1-32.
- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar, & L. A. Sroufe (Eds.), *Self-processes and development* (pp. 43-77). Hillsdale, NJ: Lawrence Erlbaum
- Dockter, D. (2010). Taking The Student's Lead In Teaching Tax. *American Journal of Business Education*, 3(7), 73-78.
- Feltham, G.D. (1996). Some thoughts on teaching tax. *CA Magazine*, 129(5), 31-33.
- Floyd, K. S., Harrington, S. J., & Santiago, J. (2009). The effect of engagement and perceived course value on deep and surface learning strategies. *Informing Science: the International Journal of an Emerging Transdiscipline*, 12, 181-190.
- Franklin, B. (1789). Benjamin Franklin Quotes. Retrieved January 12, 2017 from http://www.notable-quotes.com/f/franklin_benjamin.html
- Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, 33(4), 441-467.
- Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary educational psychology*, 29(4), 462-482.
- Handelsman, M.M., Briggs, W.L., Sullivan, N., & Towler, A. (2005). A Measure of College Student Course Engagement. *The Journal of Education Research*, 98(3), 184-191.
- Jelinek, K. (2017). Wax on, wax off: Transfer of Learning through an Experiential Learning Project. *The Accounting Educators' Journal*, 26.
- Jones, S.M., & Duncan, W.A. (1995). Teaching the introductory tax course: A new paradigm. *The Journal of the American Taxation Association*, 17(1), 95.
- Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in higher education*, 38(5), 758-773.
- Kelliher, C.F. (2014). A tax planning case using a taxpayer life-cycle approach. *Advances in Accounting Behavioral Research*, 17.
- Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, NJ, Prentice Hall
- Kolb, A. Y., & Kolb, D. A. (2009). The learning way meta-cognitive aspects of experiential learning. *Simulation & Gaming*, 40(3), 297-327.
- Komaraju, M., & Karau, S. J. (2005). The relationship between the big five personality traits and academic motivation. *Personality and individual differences*, 39(3), 557-567.

- Konak, A., Clark, T. K., & Nasereddin, M. (2014). Using Kolb's Experiential Learning Cycle to improve student learning in virtual computer laboratories. *Computers & Education*, 72, 11-22.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, 78(2), 311.
- Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education*, 79(5), 540-563.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, 31(3), 313.
- Marandos, S.A., & Randall, I. (2012). Engaging and Motivating Students: Five research based models/approaches for engaging students to be productive! *Proceedings of ASBBS*, 19(1), 539-544.
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American educational research journal*, 37(1), 153-184.
- Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology*, 106(1), 121.
- Mehrotra, A. K. (2005). Teaching Tax Stories. *Journal of Legal Education*, 55(1/2), 116-125.
- Menk, K. B., & Malone, S. (2015). Creating a Cheat-Proof Testing and Learning Environment: A Unique Testing Opportunity for Each Student. *Advances in Accounting Education: Teaching and Curriculum Innovations*, 1, 133-161.
- Petkus Jr, E. (2000). A theoretical and practical framework for service-learning in marketing: Kolb's experiential learning cycle. *Journal of Marketing Education*, 22(1), 64-70.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, 82(1), 33.
- Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. *Student perceptions in the classroom*, 7, 149-183.
- Raabe, W. (2016). Working With Today's Students. *The Tax Adviser*, 47(11), 822.
- Richmond, G. L., & Roehrenbeck, C. A. (1992). From Tedious to Trendy: A Tax Teacher's Triumph. *Nova L. Rev.*, 17, 739.
- Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: An opportunity to engage millennial students through active learning. *Journal of Family and Consumer Sciences*, 105(2), 44.
- Rubin, J.T. (1998). Teaching Tax Using a Life Cycle Approach. *The Tax Adviser*, 29(5), 334-336.
- Scogin, S. C. (2016). Identifying the Factors Leading to Success: How an Innovative Science Curriculum Cultivates Student Motivation. *Journal of Science Education and Technology*, 25(3), 375-393.

- Sigala, M., Burgoyne, C. B., & Webley, P. (1999). Short Paper Tax Communication and Social Influence: Evidence from a British Sample. *J. Community Appl. Soc. Psychol*, 9, 237-241.
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of educational psychology*, 85(4), 571.
- Sun, J. C. Y. (2014). Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. *Computers & Education*, 72, 80-89.
- Svanum, S., & Bigatti, S. M. (2009). Academic course engagement during one semester forecasts college success: Engaged students are more likely to earn a degree, do it faster, and do it better. *Journal of College Student Development*, 50(1), 120-132.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4), 295-312.
- Uniform CPA Examination – Content, Structure and Delivery FAQs . (2017). Retrieved February 26, 2017, from https://www.aicpa.org/BecomeACPA/CPAExam/ForCandidates/FAQ/Pages/computer_faqs_2.aspx
- Webber, K. L., Krylow, R. B., & Zhang, Q. (2013). Does involvement really matter? Indicators of college student success and satisfaction. *Journal of College Student Development*, 54(6), 591-611.
- Zhao, C. M., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in higher education*, 45(2), 115-138.
- Zumbrunn, S., Mckim, C., Buhs, E., & Hawley, L. R. (2014). Support, belonging, motivation, and engagement in the college classroom: A mixed method study. *Instructional Science*, 42(5), 661-684.

NAFTA/UMSCA and Ford: An Analysis of Trade Agreement Impacts on an Automotive Giant

Haleigh Casey

University of Lynchburg

caseyh586@lynchburg.edu

Eric Kyper

University of Lynchburg

kyper@lynchburg.edu

Michael Douglas

Millersville University

michael.douglas@miillersville.edu

Abstract

This paper analyzes how the North American Free Trade Agreement (NAFTA) and the United States Mexico Canada Agreement (USMCA) affect the automobile manufacturing industry, specifically Ford, with a focus on Mexico and the United States. To determine the impact realized on the company the cost of revenue data for Ford was compiled and analyzed five years pre and post each agreement. The data was analyzed using analysis of variance (ANOVA). The results indicate that NAFTA increased the cost of revenue for Ford, while USMCA had neutral results. This paper will contribute to the discussion surrounding the general efficacy of trade agreements as the enforcers of labor regulations and support mechanisms to encourage the success of specific industries.

Introduction

The research questions are whether NAFTA and USMCA contributed to lower cost of revenue for Ford. “Cost of revenue” was chosen over “costs of goods sold” because cost of revenue is more comprehensive accounting metric that includes transportation, a key component of manufacturing in Mexico for a US market.

Free trade agreements or FTAs have been crucial to economic development across the world. FTAs are legally binding agreements between two or more countries to facilitate trade by reducing or removing trade barriers. Free trade can initially disrupt certain industries causing unemployment increases, yet in the long run it decreases costs for consumers, stimulates participating economies, and facilitates technological innovation (Hudson). Both the North American Free Trade Agreement, known as NAFTA and the United States Mexico Canada Agreement, known as the USMCA, have been highly debated between politicians, economists, and consumers. One of the sectors that was impacted the heaviest by both agreements was U.S. automobile manufacturing. Initially, NAFTA dissolved the tariffs between the United States, Mexico, and Canada to facilitate efficient multilateral trade between the countries. The agreement was met in the United States with initial skepticism as it was clear that Mexico would be the primary

beneficiary. The agreement once signed allowed for U.S. manufacturers to relocate their factories in the U.S. and move across the border to use Mexico for cheaper labor. Unsurprisingly, this caused a significant increase in unemployment in the affected industries yet brought lower prices to consumers. NAFTA stood for roughly twenty years until it was renegotiated into the USMCA in 2018. This renegotiation was led by American President Donald Trump. Throughout his campaign President Trump was a strong advocate for American protectionism. He often referred to NAFTA as the ‘worst trade deal ever’ and fought to have the agreement replaced for one that has more terms favorable to the U.S. The USMCA was the improved trade deal. The new agreement stayed true to the basic principles of NAFTA yet included many new provisions, specifically regarding labor and content regulations to encourage companies to return production to the United States. Like its predecessor the USMCA is an extremely controversial agreement. Many politicians and economists are concerned about how the agreement will affect free trade amongst the participating countries in the future. This paper analyzes both agreements and their effect on the participating countries, but specifically automobile manufacturing between the United States and Mexico. The impact that NAFTA and the USMCA had on automobile manufacturing and the American workforce remains a highly debated topic, with the discussion centering around whether that effect has been positive or negative. This paper aims to answer the question of which agreement was more beneficial for the success of Ford.

Background

Free trade agreements are an integral aspect of the modern global economy. Especially in the case of the United States, Mexico and Canada. These three countries due to their geographical location are at an inherent disadvantage compared to the rest of the world. Because of the proximity, Europeans and Asian countries have an easier time maintaining global commerce. The separation from Europe and Asia by the Atlantic and Pacific Oceans was one of the main driving forces in the initial talks of NAFTA. Angeles Villarreal and Ian Fergusson dive into this history and how we got from no formal agreement to the USMCA today in both their papers, “NAFTA and the United States-Mexico-Canada Agreement (USMCA)” and “The North American Free Trade Agreement (NAFTA)” written in 2020 and 2017 respectively. In their 2017 paper they discuss how at its inception NAFTA was extremely controversial as both the United States and Canada were established developed countries, and Mexico was still classified as developing. Mexico’s journey to end their restrictive protectionist government was a main factor in their signing of the agreement, which gave them good reason to relax trade policies. However, this was a main point of contention for many Americans at the time who were concerned with the risk of significant job losses to Mexico due to their cheaper labor. To discover how NAFTA came to fruition the ratification of the Canada Free Trade Agreement (CFTA) in 1987 is essential as it was the U.S.’s first economically significant agreement. The CFTA was short-lived as it was replaced by NAFTA in 1992. NAFTA represented an attempt from the U.S. and Mexico to work together and sever any preexisting tension and apprehension about trade between the two. In “The Political Economy of NAFTA/USMCA” Gustavo Flores-Macias Mariano Sánchez-Talanquer also investigates the history of NAFTA. They agree how pre-existing CFTA mixed with Mexico’s attempts at modernization is what drove the beginning talks of NAFTA.

It is important to analyze the labor provisions of both agreements effects on the automobile manufacturing industry. In “Labor Regimes and Free Trade in North America: From the North American Free Trade Agreement to the United States–Mexico–Canada Agreement”, Francisco Campos Ortiz discusses the labor provisions of both NAFTA and the USMCA. He argues that while NAFTA itself did not contain direct labor provisions the drafting of the North American Agreement on Labor Cooperation (NAALC) as a side agreement was still significant. The side agreement was proposed as many were concerned that poor labor conditions in Mexico would benefit Mexico at the expense of the United States and their labor force rather than be mutually beneficial. The NAALC incorporated many provisions that stayed true to those upheld by the International Labor Organization (ILO). These provisions included the freedom of association, the right to collective bargaining, the right to strike, the elimination of forced labor, child labor, and employment discrimination. Additionally, the side agreement introduced minimum employment standards, equal pay for women and men, workers compensation for occupational injuries and illnesses, migrant worker protections (Campos Ortiz, 2019). In “The Governance Architecture of Transnational Labor Regulation” Patrick Witzak criticizes the NAALC. He refers to the side agreement as “a rather weak and sobering example of transnational labor regulation” because of its nationally oriented regulation (2024).

Before the signing of NAFTA many were skeptical about the effect that it would have on the economies of the participating countries. Gustavo Flores-Macias and Mariano Sánchez-Talanquer agree with the skepticism. These concerns came specifically from the United States as many economists and politicians believed that the benefit Mexico would receive from the agreement far outweighed the benefit that would be received by both the United States and Canada. These skepticisms were valid as at the time of the signing of NAFTA, “the size of the U.S. economy was 20 times that of Mexico and 10 times that of Canada. Per capita income in the United States was 8 times that of Mexico” (Flores-Macias and Sánchez-Talanquer, 2019). Additionally, the authors discuss how the wage gap between the United States and Mexico was another main point of contention for the apprehension about the agreement. This concern deals specifically with the concern that the United States labor force would end up being negatively affected as labor would be moved to where it was cheaper. It should be noted that the agreement was not only highly controversial in the United States, both Mexico and Canada also had their own reasons to be hesitant about the signing of NAFTA.

Despite all the concerns about the North American Free Trade Agreement, it was ratified in 1992. Most look to the Gross Domestic Product (GDP) as an overall indicator as it is the most widely used and accepted form to analyze a country’s overall economic performance as it encapsulates total economic growth. Both papers mention how, “NAFTA’s net impact on the U.S. economy has been relatively small, at best about 5% of total GDP” (Villarreal and Fergusson, 2017). Whether this benefits Ford has not been researched. While this marginal benefit is widely accepted, many critics of the agreement still suggest that NAFTA has come at some cost to U.S. workers, the extent to which is highly debated.

After NAFTA’s underwhelming results and growing discontent after the financial crisis in 2008 most Americans became hungry for more nationalist-protectionist policies (Flores-Macias and Sánchez-

Talanquer, 2019). This shift paired with Donald Trump's presidential campaign slogan "Make America Great Again" and his emphasis on pushing American Protectionism appealed to many American voters, and he won the election in 2016. In his inaugural address Trump reinforced the need for the United States border protection to prevent the "ravages of other countries making our products, stealing our companies, and destroying our jobs" (Campos Ortiz, 2019). After stepping into office Trump was extremely vocal about his dislike of NAFTA frequently referring to it as the "worst trade deal ever signed by the U.S" and wanted to relocate production from Mexico back to the U.S. Trump added to his attack on NAFTA by commenting on the growing trade balance between the United States and Mexico and blaming it on the agreement. The aforementioned balance began with a surplus of \$1.7 billion in 1993 to a deficit of \$63.2 billion in 2016 (Villarreal and Fergusson, 2017). Noah Zuss continued to discuss the renegotiation in his paper "Automobile Industry's Country of Origin Dispute Acting as USMCA Road Test". The U.S. auto manufacturing industry is one of the ones that drove President Trump's want to renegotiate NAFTA. During the renegotiation, Trump enacted tariffs in an attempt to prioritize U.S. produced products. This was fought strongly by Mexico and Canada because they also wanted to protect their manufacturing industries (Zuss, 2021) which made him want to replace it with what he deemed the better version, what we now know as the United States Mexico Canada Agreement (USMCA).

As this paper's purpose is to analyze the effect of both NAFTA and the USMCA on the automobile manufacturing industry, this section will discuss some effects that the agreements had on the industry as a whole. The biggest impacts for both agreements on the industry fall in their respective labor and rules of origin provisions. As mentioned earlier, the USMCA was much more comprehensive and restrictive compared to NAFTA, which is reflected in the industry performance. The USMCA increased NAFTA's rules of origin requirements for auto manufacturing, from 62.5% to 75% of components that must be produced within the NAFTA region. In addition to that requirement, 70% of each manufactured vehicle's steel and aluminum needs to originate from North America. This requirement was enforced because the governments wanted to encourage companies to keep manufacturing within the NAFTA region. This would then increase production volume, which increases employment opportunities within the region in addition to stimulating the economy. However, companies had moved production out of the region in the first place because of cheaper alternatives. This move back would mean an increase in production cost experienced by the companies, which is likely to be reflected in the prices that consumers are paying each vehicle. Robert Hudson presents an argument that companies would consider outsourcing production to other countries for production and pay tariffs as opposed to meeting the new production content requirements outlined in the USMCA. This is almost always a risk, but a barrier to this, is the fact that most automobile manufacturers that would consider relocation have persistent established facilities in the NAFTA region so the relocation costs of building new facilities and hiring new employees could outweigh the increase in costs from adhering to the USMCA provisions. Yet, companies may be willing to accept this steep price upfront if they are confident that they will be more profitable in the long run if they outsource production. In addition to the rules of origin, both agreements have labor provisions in place, again with the stricter of the two being a part of the USMCA. The wage requirement in the USMCA states that no vehicle can be manufactured at a facility where the workers are paid any less than \$16 USD per hour. This provision was enacted as an attempt by the United States to move manufacturing from Mexico back to

the United States. The initial shaft away from the U.S. was in response to NAFTA, so the U.S. attempted to rectify that with the USMCA. However, Robert Hudson notes that while the USCMA does increase the minimum wage requirement, the minimum is still \$8 below the average wage for American manufacturing workers. Because of this, while the provision was meant to be an incentive for American production, it might just force companies to increase their wages in Mexico, or as stated earlier, outsource their labor to countries where the requirement is lower or nonexistent like India or Asia (Hudson, 2021).

Overall, while the new provisions were in an attempt to revitalize the automobile manufacturing industry in the NAFTA region, it might have instead doomed it. The success of the industry and the agreement lie in the hands of the manufacturing companies. Whether they decide to absorb the new provisions, despite the cost increase or decide to outsource production will determine the outcome. This revelation is extremely ironic given that the entire reason that NAFTA was renegotiated into the USMCA was because President Donald Trump believed that NAFTA had destroyed American automobile manufacturing.

Analysis/Results

A one-way analysis of variance (ANOVA) was used to search for statistically significant differences in cost of revenue pre and post the adoptions of NAFTA/UMCSA. The data was collected using Bloomberg and covers the years 1987-1997 for NAFTA and years 2013-2023 for USMCA. All costs were inflation adjusted to a base year of 2000.

Two hypotheses were tested at the 95% confidence level ($\alpha = 0.05$).

Ho1: Ford's cost of revenue five years after the adoption of NAFTA will be the same as the five years before the adoption of NAFTA.

Ha1: Ford's cost of revenue five years after the adoption of NAFTA will be lower than the five years before the adoption of NAFTA.

Ho2: Ford's cost of revenue five years after the adoption of USMCA will be the same as the five years before the adoption of USMCA.

Ha2: Ford's cost of revenue five years after the adoption of USMCA will be lower than the five years before the adoption of USMCA.

For all results the data meet the requirements of normally distributed residuals using the Shapiro-Wilk test, and results do not change for unequal variances. Hence, the data meet the requirements of the general linear model.

NAFTA results:

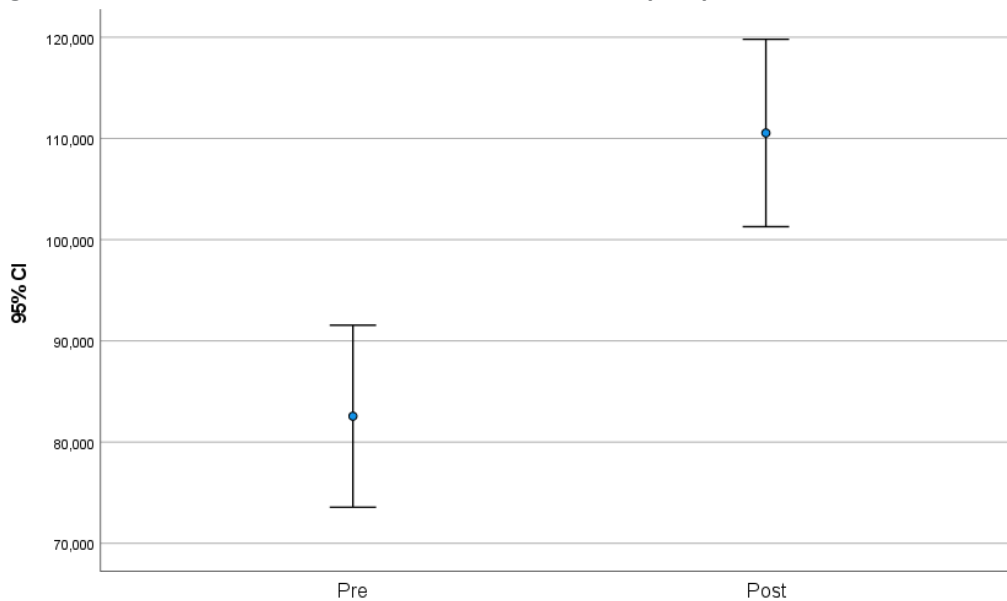
As evidenced in tables 1 and 2 and figure 1 the null hypothesis for Ho1 is rejected at the 95% confidence level. The alternative hypothesis is one-tailed; however, the results are the same for a two-tailed test.

Table 1: Group Statistics for NAFTA

Pre/Post NAFTA	n	Mean	Std. Deviation
Pre	6	111,053.33	2,539.14
Post	6	121,696.77	5,503.89

Table 2: ANOVA results for Ho1 (NAFTA)

Dependent Variable	p-statistic one-tailed test	p-statistic two-tailed test	Mean Difference
Cost Of Revenue	.005	.011	10,643.44

Figure 1: 95% confidence intervals for cost of revenue pre/post NAFTA

The results clearly show that cost of revenue post NAFTA adoption increased significantly for Ford. Some possible explanations for this will be addressed in the discussion section.

USMCA results:

Table 3: Group Statistics for USMCA (units in thousands)

Pre/Post NAFTA	n	Mean	Std. Deviation
Pre	8	133,528.5	10,200.73
Post	4	135,546.5	18,921.59

Table 4: ANOVA results for Ho2 (USMCA)

Dependent Variable	p-statistic one-tailed test	p-statistic two-tailed test	Mean Difference
Cost Of Revenue	.564	.811	2,018

The UMSCA results show that the cost of revenue pre and post policy are statistically the same. This may result from the nuanced regulations of the new agreement. However, while the cost of revenue has not increased, it has also not decreased. Since the post USMCA data is still limited there is the possibility the benefits are still accruing and have not had enough time to manifest in the data. At a minimum the new rules have reversed some of the negative effects found in the NAFTA results.

Discussion

NAFTA had a negative effect on the automobile manufacturing industry. There is a distinct negative relationship between the ratification of the agreements and the cost of revenue for Ford five years after NAFTA. There are several possible explanations for the cost of revenue increase: infrastructure costs in Mexico, transportation and logistics complications, changes in cost of supply chain in puts such as raw materials. However, what seems clear is that Ford's cost of revenue has not benefited from NAFTA. The UMSCA agreement does not appear to negatively affect Ford's cost of revenue, however it does not seem to benefit Ford either. Possibly the tax benefits of the trade agreements do not offset the costs and challenges of production in Mexico. We recognize that cost of revenue is a narrow metric, and these are complicated trade agreements to analyze in one paper. However, the expectation was that production cost would be lower with these trade agreements than without them. This paper finds no evidence to support that assumption. Trade agreements have shown to have an effect on the automobile industry, so it is possible that with revisions free trade agreements have the ability to support the automotive industry. Those revisions would likely need to be tailored to the needs of a specific industry and/or company like Ford.

Future Research and Limitations

Future research should focus on other trade agreements and how they affect various industries. After observing the negative impact these two agreements had on America's largest automobile manufacturer, it will of value to see if this effect would be different if certain provisions were changed. That also brings the question of whether it is possible to construct an agreement that does not have a negative impact on at least one industry. Future research should also explore how free trade agreements have affected industries and how they can be structured to either destabilize or support different industries through provisions and side agreements. In addition, as modern free trade agreements are being used as ways to enforce labor provisions, more research is needed to determine the efficacy of using agreements as an enforcement tool. As the USMCA contains the most comprehensive labor provisions, future agreements and research should use it as a model. Whether the USMCA is ultimately successful will show in time, but whether it succeeds or fails it will serve as a blueprint for future agreements inclusion and detail of labor provisions.

As noted earlier, it is not possible to say for certain that the increase in cost of revenue Ford in the observed period came exclusively from the signed trade agreements. Most economic indicators are interrelated to the point where the tracking of specific cause and effect relationships can be difficult, maybe impossible to ascertain. The results indicate the existence of a statistically significant relationship

between the agreements and cost data, yet there always exists the possibility of error, even if the likelihood of it is extremely minute. Furthermore, there could have been other variables that caused this increase that were not incorporated into this analysis.

Finally, we did want to explore the same questions for General Motors. However, the reports for General Motors were more difficult to retrieve. This is because after the Great Recession in 2008, GM filed for Chapter 11 bankruptcy and reorganized their company. This reorganization involved the company changing who they do their financial reporting through. Because of the legal aspects of their bankruptcy, the financial information for the aforementioned years is not publicly available.

Conclusion

After conducting a comparative analysis of both the North American Free Trade Agreement and the United States Mexico Canada Agreement and their impact on the automobile manufacturing industry this paper concludes that Ford's cost of revenue was negatively impacted by both agreements.

Citations

Campos Ortiz, Francisco E. "Labor Regimes and Free Trade in North America: From the North American Free Trade Agreement to the United States–Mexico–Canada Agreement." *Latin American Policy*, vol. 10, no. 2, Nov. 2019, pp. 268–285, <https://doi.org/10.1111/lamp.12171>.

Flores-Macias, Gustavo A., and Mariano Sánchez-Talanquer. "The Political Economy of NAFTA/USMCA." *Oxford Research Encyclopedia of Politics* (2019).

Hudson, Robert "The USMCA: An Ideal 'New NAFTA?'" Ky. L. J. Online (April 24, 2021), <https://perma.cc/P6G9-Y7N2>.

Villareal, Angeles M., and Ian F. Fergusson. "The North American Free Trade Agreement (NAFTA)." *Congressional Research Service Report* (2017).

Villarreal, Angeles M., and Ian F. Fergusson. "NAFTA and the United States-Mexico-Canada Agreement (USMCA)." *Congressional Research Service Report* (2020).

Witzak, Patrick. "The Governance Architecture of Transnational Labor Regulation." *Routledge, Labor History*, 17 July. 2024, DOI: 10.1080/0023656X.2024.2380849

Zuss, Noah. "Automobile Industry's Country of Origin Dispute Acting as USMCA Road Test." *International Financial Law Review*, 2021. ProQuest, <https://lynchburg.idm.oclc.org/login?url=https://www.proquest.com/trade-journals/automobile-industry-s-country-origin-dispute/docview/2589716210/se-2>.

Zuss, Noah. "USMCA could Solve Supply Chain Problems." *International Financial Law Review*, 2021. *ProQuest*, <https://lynchburg.idm.oclc.org/login?url=https://www.proquest.com/trade-journals/usmca-could-solve-supply-chain-problems/docview/2606753545/se-2>.

Paradoxes in Combining Inequalities

Jinchang Wang

School of Business, Stockton University

Galloway, New Jersey 08215

jinchang.wang@stockton.edu

Abstract

We discuss the paradoxes that occur when adding inequalities or equations, in which legal algebraic operations lead to false or absurd outcomes. Diagnostic analyses indicate that perplexities are caused by our misunderstanding of the operation of combining inequalities in algebra and set theory. Inequality in algebra compares the size of two numbers. While inequality in set theory represents a set. An algebraic operation is not necessarily legal in non-algebra circumstances such as analytical geometry, linear programming, and set theory. Adding two inequalities up is not a universally valid operation. Adding two inequalities maintains the relationship in algebra, but it does not maintain the set represented by the inequalities. The paradoxes remind us of a deep-rooted misconception in dealing with simultaneous inequalities. A guideline of dealing with simultaneous equations and inequalities in applications is provided.

Key Words: Inequality, Equation, Algebra, Set Theory, Linear programming, Data Analytics, Analytical Geometry, Logic.

A *paradox* refers to bewilderment in which apparently sound reasoning and legal operations leads to a conclusion that seems senseless or absurd.

U. Examples of Paradoxes in Combining Inequalities

We have taken it for granted since early ages that algebraic rules about inequalities are valid as the sun rises from the east. Examples below show bizarre results by adding two equations or two inequalities up.

I-1. Paradox 1:

Given a system of simultaneous equations:

$$X^2 - 4 = 0, \quad (1.1)$$

$$X + 2 = 0. \quad (1.2)$$

Combining the two equations by using Rule-A,

$$(X^2 - 4) + (X + 2) = 0 + 0, \text{ i.e.,} \\ X^2 + X - 2 = 0. \quad (1.3)$$

Solutions to equation (1.3) are:

$$X=-2 \text{ and } X=1.$$

Note:

Simultaneous equations (1.1) and (1.2) have one solution which is $X=-2$. After combining them, the resulting equation (1.3) has two solutions, $X=-2$ and $X=1$, in which $X=1$ is not a solution to the original system of simultaneous equations (1.1) and (1.2) at all!

I-2. Paradox 2:

Given a system of simultaneous inequalities:

$$X > 2, \quad (2.1)$$

$$X > 3. \quad (2.2)$$

Add (2.1) to (2.2) by Rule-B:

$$2X > 5 \quad (2.3)$$

After simplifying:

$$X > 2.5. \quad (2.4)$$

Note:

Solution to simultaneous inequalities (2.1) and (2.2) is $X > 3$. But after adding two inequalities up, we have $X > 2.5$ as in (2.4), which is not the solution to simultaneous inequalities (2.1) and (2.2)! For example, $X=2.8$ satisfies (2.4), but does not satisfy both (2.1) and (2.2).

I-3. Paradox 3:

Given:

$$5 \geq X \geq 2. \quad (3.1)$$

Expressed (3.1) with two inequalities:

$$5 \geq X, \quad (3.2)$$

$$X \geq 2. \quad (3.3)$$

Multiply (3.3) by (-1) , and flip the resulted inequality up:

$$-2 \geq -X. \quad (3.4)$$

Add (3.2) and (3.4) by using Rule-B:

$$5 - 2 \geq X - X \quad (3.5)$$

After simplifying:

$$3 \geq 0. \quad (3.6)$$

Note:

Inequality (3.6) is true all the time (tautology). But (3.6) does not contain the restrictive information that X is between 2 and 5 as given in (3.1)! The information is lost in the operation of adding (3.2) and (3.4).

I-4. Paradox 4

This is an example in integer programming.

Given a proposition:

If A is on then either B or C is on, but not both; if A is off then B and C can be either on or off. (4.1)

Three binary variables are defined:

$X_a=1$ if A is on, $X_a=0$ if A is off.

$X_b=1$ if B is on, $X_b=0$ if B is off.

$X_c=1$ if C is on, $X_c=0$ if C is off.

Proposition (4.1) is represented by inequalities with binary variables X_a , X_b , and X_c :

$$X_a \leq X_b + X_c, \quad (4.2)$$

$$X_b + X_c \leq 1. \quad (4.3)$$

Add (4.2) and (4.3) by using Rule-B:

$$X_a + X_b + X_c \leq X_b + X_c + 1. \quad (4.4)$$

After simplifying:

$$X_a \leq 1. \quad (4.5)$$

Note:

Simultaneous inequalities (4.2) and (4.3) correctly represent proposition (4.1) with binary variables. After a “legal” operation by Rule-B, we get $X_a \leq 1$ as in (4.5), which says X_a can be either 0 or 1. The original restrictive information about B and C is missing in (4.5)!

I-5. Paradox 5.

In this paradox, inconsistent constraints of a linear program become consistent after combining constraints.

Given constraints of a linear program:

$$X + 0.5Y \leq 1, \quad (5.1)$$

$$2X + Y \geq 3, \quad (5.2)$$

Multiply (5.1) by (-1):

$$-X - 0.5Y \geq -1. \quad (5.3)$$

Add (5.2) and (5.3) up by Rule-B:

$$X + 0.5Y \geq 2 \quad (5.4)$$

Note:

The original feasible solution region of the linear program is empty because (5.1) and (5.2) are inconsistent (can be easily verified).

However, the feasible solution region, represented by (5.4), after adding (5.1) and (5.2) up is obviously non-empty.

Adding two inequalities up could make an empty feasible region non-empty!

I-6. Paradox 6.

This paradox demonstrates that combining constraints could result in looser restrictions.

Given constraints of a linear program:

$$X + 2Y \leq 6, \quad (6.1)$$

$$2X+Y \leq 6, \quad (6.2)$$

Adding (6.1) and (6.2) by using Rule-B:

$$3X+3Y \leq 12. \quad (6.3)$$

After simplification:

$$X+Y \leq 4. \quad (6.4)$$

Note:

The feasible region defined by (6.4) is different from the original feasible region defined by (6.1) and (6.2). For example, solution ($X=0, Y=4$) is not feasible according to (6.1) and (6.2), but it is feasible according to (6.4) which is the sum of (6.1) and (6.2)!

II. What Causes the Paradoxes

The above six examples of paradoxes alarm us that there is something wrong, or wrongly understood, somewhere in fields of algebra, analytical geometry, linear programming, integer programming, or set theory. Let us make a pathological analysis.

To put our pathology short, the perplexity of the above paradoxes is caused by our misunderstanding and misconception about operations of inequalities. Adding equations or adding inequalities is not valid among “sets” in set theory and its applications, although it is always valid among “numbers” in algebra,.

Inequality $X \geq A$ in algebra tells the relationship between two numbers, X and A . For example, $X \geq 5$ tells the number represented by X is greater than or equal to 5.

Inequality $X \geq A$ in set theory represents a set of numbers each of which is greater than or equal to A . For example, $X \geq 5$ represents a set in which every element is greater than or equal to 5.

Adding two inequalities is a legal operation in algebra; but it is not valid in set theory. Adding two inequalities of same direction is a legal operation in algebra, because it maintains the relationship of inequality. But adding two inequalities does not maintain the set represented by the two inequalities. In fact, between two sets, union $\cup$ and intersection $\cap$ are legal operations, while “adding two inequalities representing the sets” is not valid at all.

Most of us started to have this misconception in middle school. We falsely think all we have learned about equations and inequalities in algebra are universally valid everywhere. We were not warned by the textbooks about the danger of adding inequalities in linear programming, integer programming, and analytical geometry.

Algebra is investigating property, relation, and calculations of individual numbers, while set theory is investigating property, relation, and calculations of sets which are groups of numbers. Such a difference between algebra and set theory is not well distinguished in our algebra courses and textbooks.

III. Comments and Discussions:

We have diagnosed and clarified the perplexities ascribable to adding inequalities. So, the paradoxes listed in Section I are no longer paradoxes now. Considering that the misconceptions on combining inequalities are deep-rooted in people's mind, we make following guidelines and suggestions for dealing with inequalities.

- (U) Never combine or add inequalities in the following circumstances where inequalities are used to represent a set of values:
Solving a system of simultaneous inequalities,
Analytical geometry,
Linear programming,
Integer programming,
Non-linear programming.

If one does want to combine multiple inequalities for the purpose of simplification, for example in integer programming's 0-1 constraints, then make sure to verify the result after the combination to make sure it is equivalent to the original system.

(2) Use caution in solving simultaneous equations. An equation is a special case of inequality. Solving simultaneous inequalities is a problem about "sets", so is solving simultaneous equations. When we combine equations in the process of solving simultaneous equations, we must bear in mind that the solution to the original system is a subset of the derived solutions, as shown in the example of Paradox 1 in section I-1.

(3) The issue of reasoning in a knowledge base or database in AI can be converted to a problem of verifying consistency of a system of simultaneous inequalities [Jeroslow and Wang 1990], which is a computationally hard NP-complete problem. Never combining inequalities for the purpose of simplification, since inequalities represent "set" in this circumstance.

Acknowledgements:

I appreciate Mr. Paul Kline, a student in my MBA class “Managerial Decision Analysis”, who raised the question as shown in Paradox 4 in I-4. His question is the initial impetus for the research presented here.

I appreciate Professor Xiaojun Shen and Ms. Jinyuan Wang for their enlightening thoughts and analyses in discussions of some paradoxes addressed in this article.

References:

[Jeroslow and Wang] Jeroslow, Robert G. and Jinchang Wang, Solving Propositional Satisfiability Problems, *Annals of Mathematics and Artificial Intelligence*, 1, 1990, 167-187

Factors that Affect American Savings in the 21st Century

Grace Gallagher

Zarb School of Business, Hofstra University

Hempstead, New York, 11549

ggallagher1@pride.hofstra.edu

ABSTRACT

Savings behavior is a well-explored topic, but there is a gap in the literature regarding the factors influencing American savings behavior in the 21st century. This paper aims to fill that gap by examining the factors that affect savings behavior across and within categories, with a specific focus on those with less than \$100k saved. Using logistic and beta regressions, we analyze both categorical and continuous predictors to identify key determinants of savings. Significant predictors in both regressions include income level, homeownership, retirement savings accounts, investments outside retirement accounts, emergency expenses and credit card usage, education level, and self-employment. The effects of these predictors vary between the two models. Some predictors, such as age and student loan debt, are significant only in the logistic regression, while race/ethnicity is significant only in the beta regression. Surprisingly, gender, marital status, and the presence of children are insignificant in both models, despite being significant in previous studies. These findings provide more nuanced perspectives on American savings behavior.

Key Words: *savings behavior, logistic regression, beta regression, American, 21st century*

U. INTRODUCTION

The ability to save is an essential pillar of financial security. However, a substantial number of Americans struggle to save adequately. Forty percent of all Americans cannot plan beyond their next paycheck, and 26 percent do not have \$500 set aside for emergencies (New National Poll: Economic Hardships of Millions of Middle-Class Americans Go Unseen; *Most Believe They Will Continue to Struggle Throughout Their Lives*, 2024).

Previous research suggests that a range of factors—from income levels to access to financial services—play a significant role in shaping an individual's saving habits. This study investigates factors contributing to American savings behavior specifically in the 21st century and determines if previous literature applies to America's current savings behavior. Understanding potential causes of saving discrepancies is essential for designing interventions that help individuals achieve greater financial security.

2. LITERATURE REVIEW

Many studies explored factors impacting savings behavior, focusing on different countries and periods. Some emphasize easily measurable predictors like income and age, while others investigate abstract concepts like financial literacy. Given the quantitative nature of our data, we examined the effect of measurable predictors on savings behavior.

Most previous literature mentions income as a significant predictor of savings. Shin and Kim (2018) found income positively correlates with savings. Similarly, Callen and Thimann (1997) and Hefferan (1982) found a positive link between income growth and savings. Meanwhile, Davis and

Schumm (1987) found income significant for savings only among Americans earning above \$9k annually. These papers all highlight the importance of income in understanding human saving behavior across various periods and geographic locations, even if their conclusions are slightly different from one another.

Heckman and Hanna (2015) found that low-income Americans with employer-sponsored retirement plans are 1.17 times more likely to save. This is the only piece of literature to mention employer-sponsored retirement plans as a significant predictor of savings behavior. Additionally, Heckman and Hanna (2015) found education insignificant in predicting savings. Conversely, Davis and Schumm (1987) and Yuh and Hanna (2010) found higher education levels associated with higher savings rates. These findings suggest that both the presence of employer-sponsored retirement plans and an individual's education level may serve as important predictors of savings behavior among Americans.

Various pieces of literature mention homeownership, however its significance in predicting savings behavior is inconsistent across all studies. Hogarth and Anguelov (2003) found homeownership not significant for low-income savings, while Yuh and Hanna (2010) and Hefferan (1982) found positive correlations between homeownership and saving. Previous studies mentioned age as a predictor of savings behavior. However, the effect of age on savings behavior varies between studies. Shin and Kim (2018) found savings increase until age 56, after which dissaving occurs. Joseph (2019) also found that age plays a strong role in American savings behaviors. However, Hogarth and Anguelov (2003) found age insignificant for low-income Americans. These are additional predictors that may aid in understanding American saving behaviors in the 21st century.

Researchers also investigated gender and marital status relative to savings behavior. Khandu et al. (2021) found no gender difference in savings among students in Bhutan, while Furnham (1985) identified gender as a significant predictor. Regarding marital status, Heckman and Hanna (2015) found singles less likely to save than married couples. Yuh and Hanna (2010) confirmed single females save less, while Hogarth and Anguelov (2003) found marital status insignificant. While the findings regarding gender and marital status are not entirely consistent, the inclusion of these variables remains important. They may contribute to a more comprehensive understanding of American savings behavior in the 21st century or help identify factors that are not critical in explaining American savings patterns.

Several previous pieces of literature investigated the presence of children. Heckman and Hanna (2015) found families with children under eighteen are less likely to save. Hogarth and Anguelov (2003) echo that sentiment as they found families with younger children are also less likely to save than those with older children. One's ethnicity and race also appeared in two previous pieces of literature as a possible predictor of savings behavior. Heckman and Hanna (2015) found no racial/ethnic differences in savings, while Joseph (2019) found ethnic background relates to saving behavior. The presence of children appears to be an important predictor of savings behavior in the U.S., as suggested by previous studies. In contrast, the role of race and ethnicity in savings behavior seems to be less pronounced, though still a potential factor to consider. Both variables, however, remain relevant in understanding American savings behavior.

Finally, previous literature studied one's employment status and access to a bank account as predictors. Ramanathan (1969) found self-employed individuals save more, while Yuh and Hanna (2010) found retirees less likely to save. Callen and Thimann (1997) found unemployment detrimental to savings. Hogarth and Anguelov (2003) found households with bank accounts 1.8 times more likely to save. Given these findings, employment status and access to a bank account are considered relevant variables for understanding American savings behavior.

Based on this literature, we aim to identify key factors impacting savings behavior among Americans in the 21st century. We state our goal in the following research question:

RQ: What factors affect saving behaviors in Americans in the 21st Century?

3. METHODOLOGY

3.1. Data

The Survey of Household Economics and Decision Making, conducted annually by the Federal Reserve Board since 2013, examines the financial well-being of U.S. households. The 2023 survey included over 11,400 American adults and covered topics such as savings habits, credit access, student loans, and retirement planning. This study captures total household savings and investments, categorized as: Not Sure, Under \$50,000, \$50,000-\$99,999, \$100,000-\$249,999, \$250,000-\$499,999, \$500,000-\$999,999, and \$1,000,000 or more. We excluded 1,839 respondents who did not answer this question from the analysis.

3.2 Logistic Regression

As previously stated, this paper focuses on individuals with less than \$100,000 in savings. We created a binary variable for this group, with a value of 1 indicating an individual with less than \$100k saved, and 0 indicating people with more than \$100k saved. We then performed a logistic regression to identify significant predictors of being in the ‘less than \$100k’ savings category compared to the ‘more than \$100k’ savings category. We refined the model to include only predictors that were significant at the .001 level. Although several predictors were significant at the .05 and .01 levels, they did not substantially improve the model’s accuracy and thus did not warrant inclusion. Table 1 presents the results, including estimates and p-values for each predictor.

Table 1: Logistic Regression Comparing Those with < \$100k Saved to Those with > \$100k Saved.

Variable	Estimate	P-value
Ppinc7 (Less than \$10,000)		
\$10,000 to \$24,999	.455508	.030515
\$25,000 to \$49,999	-.296237	.112468
\$50,000 to \$74,999	-.702385	0.000172
\$75,000 to \$99,999	-1.095431	7.61e-09
\$100,000 to \$149,999	-1.696579	<2e-16
\$150,000 or more	-2.284952	<2e-16
Ppage	-0.016745	<2e-16
GH1 (Neither own nor pay rent)		
Own your home free and clear (without a mortgage or loan)	0.544004	1.22e-05
Own your home with a mortgage or loan	1.125113	<2e-16
Pay rent	1.408397	<2e-16
SL1Yes	.660300	1.16e-12
K21_aYes	-0.771636	<2e-16
K21_cYes	-0.674608	<2e-16
IO_bYes	-0.538750	<2e-16
EF3_bPut it on my credit card and pay it off over time	0.506220	2.36e-10
BK1Yes	0.806887	7.16e-08
ED0 (Less than high school degree)		
High school degree or GED	-0.344353	0.029919

Some college but no degree (including currently enrolled in college)	-0.497815	0.002055
Certificate or technical degree	-0.343987	0.075183
Associate degree	-0.527813	0.002416
Bachelor's degree	-0.527965	0.001400
Master's degree	-0.580597	0.000984
Professional degree (e.g., MBA, MD, JD)	-0.821373	9.63e-05
Doctoral Degree	-0.599935	0.011889
Ppfs1482 (Don't know)		
Very poor	1.436660	1.98e-09
Poor	1.693797	5.20e-16
Fair	1.318321	<2e-16
Good	1.172407	<2e-16
Excellent	0.489002	4.91e-05
Ppcm13011 (Government)		
Non-profit organizations include tax-exempt and charitable organizations	-0.15162	0.222537
Not asked	-0.754151	5.90e-14
Private-for-profit company	-0.154344	0.088978
Refused	-0.65783	0.410428
Self-employed	-0.45498	0.000363
Working in the family business	-0.171207	0.594767

(p-values less than 0.001)

These predictors accurately guess whether someone has less than \$100k in savings 79.4% of the time. Below are the interpretations for each of the variables in our final model.

Income (ppinc7) is a significant predictor in understanding the likelihood of Americans having less than \$100k in savings. The odds of having less than \$100k saved for individuals earning \$50,000 - \$74,999 is about 50.4% lower than those earning less than \$10,000. The odds of having less than \$100k saved for individuals earning \$75,000 - \$99,999 is about 66.6% lower. Additionally, the odds of having less than \$100k saved for individuals earning \$100,000 - \$149,999 is about 81.7% lower. Finally, the odds of having less than \$100k saved for individuals earning \$150,000 or more are about 89.8% lower than those earning less than \$10,000. Our analysis also tells us that for individuals earning \$10,000 - \$24,999 and \$25,000 - \$49,999, there is no statistically significant difference in the odds of having less than \$100k saved compared to those who make less than \$10,000.

Age (ppage), as well as homeownership (GH1), are also significant in understanding the likelihood of Americans having less than \$100k saved. As a person's age increases by one year, their odds of having less than \$100,000 in savings decrease by 1.66%. Additionally, individuals who own their home free and clear have 1.72 times the odds of having less than \$100k saved compared to individuals who neither own nor rent. People who own their home with a mortgage have 3.08 times the odds of having less than \$100k saved, while those who pay rent have 4.09 times the odds of having less than \$100k in savings compared to individuals who neither own nor rent.

Student loan debt (SL1), though not mentioned in the above literature, is another predictor that our model deems significant in understanding the savings behavior of Americans. Individuals currently with student loan debt from their education have 2.24 times the odds of having less than \$100k saved compared to individuals without student loan debt.

The presence of a retirement savings account (K21_a), access to stocks, bonds, etc. (K21_c), and diversified forms of income (I0_b) are also significant in understanding the likelihood of Americans having less than \$100k saved. The odds of Americans with a retirement savings account, such as a 401(k) plan through an employer, IRA, or Roth IRA having less than \$100,000 saved are 53.8% lower than those without such an account. Additionally, the odds of having less than \$100k saved for people with stocks, bonds, ETFs, or mutual funds held outside a retirement account are 49.1% lower than those without these resources. Finally, the odds of having less than \$100k in savings for people who received income from interest, dividends, or rental income in the past 12 months are 41.7% lower than those without such income streams.

How an individual deals with an emergency expense (EF3_b), and whether they have access to a bank account (BK1) is also indicative of their likelihood of having less than \$100k saved. Those who put a \$400 emergency expense on their credit card and pay it off over time have 1.66 times the odds of having less than \$100k saved compared to individuals who would not manage the emergency expense this way. Also, those with a checking, savings, or money market account have 2.24 times the odds of having less than \$100,000 in savings relative to those without such an account.

Education level (ED0) is important in understanding Americans' likelihood of having less than \$100k saved, as discussed by previous literature. The odds of having less than \$100k saved for individuals with a master's degree are about 44.0% lower compared to those with less than a high school degree. The odds of having less than \$100k saved for people with a professional degree are about 56.0% lower relative to those with less than a high school degree. While the odds for individuals with a Doctoral degree are not statistically significant at the 0.001 level, they are close to being significant at the 0.01 level. Therefore, at the 0.01 significance level, the odds of having less than \$100k in savings for individuals with a Doctoral degree are approximately 45.1% lower compared to those with less than a high school degree. Additionally, for individuals with a high school degree or GED, some college but no degree (including those currently enrolled in college), a certificate or technical degree, an associate degree, or a Bachelor's degree, there is no statistically significant difference in the odds of having less than \$100k saved compared to those with less than a high school degree.

Credit card score (ppfs1482) is another significant predictor that helps understand one's likelihood of having less than \$100k saved that was not mentioned in previous literature. Individuals who think their credit card score is very poor have 4.21 times the odds of having less than \$100k saved compared to individuals who do not know where their credit card score falls. Individuals who think their score is poor have 5.44 times the odds, people who think their score is fair have 3.74 times the odds, people who think their score is good have 3.23 times the odds, and people who think their score is excellent have 1.63 times the odds of having less than \$100k in savings compared to individuals who do not know where their score lies.

Self-employment (ppcm13011) is the final variable significant to this model in understanding American savings behaviors. The odds of having less than \$100k saved for self-employed individuals are 36.6% lower compared to people who work for the government. The odds of having less than \$100k saved for those working at non-profit organizations including tax-exempt and charitable organizations, a private-for-profit company, those who refused to answer, or those working in the family business are statistically the same as those who work for the government.

This logistic regression allowed us to identify the statistically significant predictors of American savings behavior, specifically concerning the categorical outcome of whether an individual has saved more than or less than \$100k. This provided valuable insight into the variables that contribute to the likelihood of having less than \$100k in savings. Some of the predictors overlap with prior literature, while some give us a new lens through which we can understand American saving behavior across categories.

3.3 Beta Regression

We also sought to examine these predictors using continuous savings data by attributing a specific savings amount to everyone. This approach allowed us to assess whether the significant predictors identified in the logistic regression remained consistent and to explore their effects specifically within the group of individuals who have less than \$100k in savings.

The original dataset lacked specific savings amounts. Thus, we simulated savings amounts for 8,126 observations across six categories.

Savings Category	Number of Observations
Under \$50,000	3,896
\$50,000 - \$99,999	1,163
\$100,000 - \$249,999	1,323
\$250,000 - \$499,999	922
\$500,000 - \$999,999	822

For each savings category, we generated values using a random uniform distribution in R, excluding the ‘Over \$1,000,000’ category due to its uncapped nature. After assigning these simulated values to the corresponding observations, we found that the data followed a beta distribution, which justified the use of beta regression. As the next step, we performed a beta regression with a logit link to examine the effect of each predictor on the savings behavior of Americans with less than \$100k saved.

Table 2: Beta Regression for those with less than \$100,000 in Savings.

Variable	Estimate	P-value
Ppinc7 (Less than \$10,000)		
\$10,000 to \$24,999	-0.127336	0.001242
\$25,000 to \$49,999	-0.112082	7.70e-05
\$50,000 to \$74,999	-0.056775	0.006774
\$75,000 to \$99,999	0.069002	0.001121
\$100,000 to \$149,999	0.057006	0.017998
\$150,000 or more	0.079530	0.001461
K21_aYes	-0.040063	0.000987
K21_bYes	-0.052848	1.50e-05
K21_cYes	-0.062078	3.42e-06
K21_dYes	-0.023443	0.034824
I0_bYes	-0.026317	0.049366
EF1Yes	-0.065430	1.60e-08
EF3_bPut it on my credit card and pay it off over time	0.046206	0.000206
GH1 (Neither own nor pay rent)		
Own your home free and clear (without a mortgage or loan)	0.054738	0.053187
Own your home with a mortgage or loan	0.049108	0.020611
Pay rent	-0.024227	0.163199
Ppeduc5 (no High school diploma or GED)		
High school graduate (high school diploma or the equivalent GED)	0.014636	0.622074
Some college or associate’s degree	-0.039957	0.031475
Bachelor’s degree	0.005661	0.746492
Master’s degree or higher	-0.001225	0.954497

Ppcm13011 (Government)		
Non-profit organizations include tax-exempt and charitable organizations	-0.135583	0.003273
Not asked	-0.020920	0.663595
Private-for-profit company	-0.031862	0.9444874
Refused	-0.104771	0.010992
Self-employed	0.097538	0.645830
Working in the family business	0.044569	0.352145
Ppfs1482 (Don't know)		
Very poor	0.073517	0.028616
Poor	-0.002101	0.96080
Fair	-0.085137	0.011821
Good	-0.019626	0.426399
Excellent	0.010815	0.601443
Race_5cat (White)		
Asian	0.013155	0.487275
Other	0.007189	0.876711
Hispanic	0.027477	0.292583
Black	-0.061547	0.015710

(p-values less than 0.05)

These interpretations apply only to individuals who saved less than \$100,000. This time, we kept all variables significant at the .05 level in our model.

Income (ppinc7) is also a significant predictor in this model, helping us understand the effect of income on the savings behavior of Americans with between \$0 and \$100k saved. The odds of having a higher savings amount (closer to \$100,000) for individuals earning \$10,000 to \$24,999 are 12.0% lower than those earning less than \$10,000, based on an odds ratio of 0.880. The odds for individuals earning \$25,000 to \$49,999 are 10.6% lower, and the odds for those earning \$50,000 to \$74,999 are 5.5% lower compared to individuals earning less than \$10,000. Conversely, individuals earning \$75,000 to \$99,999 have 1.07 times the odds of having a higher savings amount (closer to \$100,000) than those earning less than \$10,000. Individuals earning \$100,000 to \$149,999 have 1.06 times the odds, and those earning \$150,000 or more have 1.08 times the odds of having a savings amount closer to \$100,000 compared to those earning less than \$10,000.

The presence of a retirement savings account (K21_a) is also significant to this model. The odds of having closer to \$100k saved for those with a retirement savings account, such as a 401(k) plan through an employer, IRA, or Roth IRA are about 3.9% lower compared to those without such resources. Interestingly, the presence of a pension (K21_b) is significant in understanding the savings behavior of Americans with \$0 - \$100k saved, but not significant in the previous logistic regression. The odds of having more saved (closer to \$100k) for those with a pension with a defined benefit through an employer that will pay a monthly amount in retirement are about 5.1% lower relative to those without such a pension.

Additionally, the odds of having more saved for those with stocks, bonds, ETFs, or mutual funds held outside a retirement (K21_c) are about 6.0% lower relative to those without such resources. This is another variable significant in both regressions. The odds of having more savings for those with a savings account, money market account, or certificate of deposit (CD) (K21_d variable) are 2.3% lower compared to those without these resources.

Having diversified forms of income (IO_b) is also significant in the logistic and beta regressions. The odds of having closer to \$100k saved for those who received any income from interest, dividends, or rental income in the past 12 months are 2.6% lower relative to those without this.

The presence of an emergency fund (EF1) is significant in understanding the savings behaviors of Americans with \$0 - \$100k saved. This predictor's significance aligns with the literature from above. The odds of having higher savings for those with emergency or rainy-day funds that would cover expenses for 3 months in case of sickness, job loss, economic downturn, or other emergencies are 6.3% lower compared to those without such a fund.

Individuals who put an emergency expense that costs \$400 on their credit card and pay it off over time (EF3_b) have 1.05 times the odds of having closer to \$100k saved than those who would not do this. This variable is significant in understanding American savings behavior across, as well as within specific savings groups.

Homeownership (GH1) and education (ppeduc5) are more examples of predictors significant in both our models. Individuals who own their home with a mortgage or loan are 1.05 times more likely to have closer to \$100,000 in savings compared to neither own nor rent. Considering people who own their home free and clear without a mortgage or those who rent, there is no statistically significant difference in the odds of having closer to \$100k saved compared to those who do not rent nor own.

The odds of having closer to \$100k saved for people with some college or associate's degree are 4.5% lower compared to people with no high school degree. Additionally, our analysis tells us that for individuals with a high school diploma or the equivalent GED, a bachelor's degree, or a master's degree or higher, there is no statistically significant difference in the odds of having closer to \$100k saved compared to those who do not have a high school diploma or GED.

Employment status (ppcm13011), as well as perceived credit card score (ppcm1482), are useful in understanding the savings behaviors of Americans with \$0 - \$100k saved. The odds of having more saved (closer to \$100k) for people working for a non-profit organization including tax-exempt and charitable organizations are 12.6% lower compared to individuals who work in government. There is no statistically significant difference in the odds of savings between those who are self-employed, those who work for a private-for-profit company, and those who work in the family business versus those working a government job.

Individuals who think their credit card score is very poor have 1.08 times the odds of having closer to \$100k saved compared to individuals who do not know where their credit card score falls. The odds of having closer to \$100k saved for individuals who think their score is fair are 8.2% lower compared to those who do not know where their score falls. Additionally, there is no statistically significant difference in the odds of savings between those who think their score is either poor, good, or excellent versus those who do not know where their score falls.

Finally, race/ethnicity (race_5cat) is significant in understanding the savings behaviors of Americans with less than \$100k saved. The odds of having closer to \$100k saved for people that fall into the 'Other' ethnicity category are 6.0% lower compared to White people. The odds of having closer to \$100k saved are statistically the same when comparing Hispanic or Black people to White individuals.

This beta regression provides new insights into the effects of our predictors on the savings behaviors of Americans who have less than \$100k saved. It highlights which predictors remain significant across both regressions and examines how these predictors impact savings, whether in terms of savings categories or numeric savings amounts. Additionally, the regression reveals which predictors offer a clearer understanding of savings behavior when considering specific numeric savings amounts, as opposed to when comparing savings categories.

4. DISCUSSION

Taking the findings of both regressions into account, we will now analyze the variables that are significant in both models and their relation to the above literature.

1. Income Level.

When comparing savings categories, individuals who earn \$50k - \$74.9k, \$75k - \$99.9k, \$100k - \$149.9k, or \$150k or more have increasingly lower odds of having less than \$100k saved compared to those who earn less than \$10k. This finding aligns with Callen and Thimann (1997) and Hefferan (1982), who emphasize the positive relationship between income and savings behavior. Among individuals who have less than \$100k saved, those earning \$75k - \$99.9k, \$100k - \$149.9k, or \$150k or more have higher odds of having closer to \$100k in savings compared to those earning under \$10k. However, the odds of saving closer to \$100k are lower for individuals earning between \$10k - \$24.9k or \$25k - \$49.9k compared to those making less than \$10k. These findings suggest that the beta regression results partially support the conclusions of Shin and Kim (2018), but also partially contradict their conclusions. The effect of income on savings depends on whether savings is examined across or within groups.

2. Home Ownership.

Homeownership is a significant factor when comparing savings categories, as homeowners (whether mortgage-free or with a mortgage) have 1.72 and 3.08 times the odds, respectively, of having less than \$100k saved compared to those who neither own nor rent. Renters are 4.09 times more likely to have less than \$100k saved. This finding contradicts Yuh and Hanna (2010), who found homeowners to be more likely to save. When analyzing only those with less than \$100k saved, individuals who own a home with a mortgage are 1.05 times more likely to have savings closer to \$100k. However, individuals who own their home free and clear, as well as renters, have statistically the same odds of having closer to \$100k saved compared to those who neither own nor rent. Thus, the beta regression results partially align with Yuh and Hanna (2010) and Hefferan (1982), but also partially contradict their conclusions in the context of individuals who have less than \$100k saved.

3. Retirement Savings Accounts.

Comparing savings groups, the odds of having less than \$100k saved for individuals with retirement savings accounts (e.g., 401(k), IRA) are 53.8% lower than those without such accounts. This finding aligns with Heckman and Hanna (2015), who found that individuals with retirement accounts are more likely to save. However, when considering only those with less than \$100k saved, individuals with retirement savings accounts have 3.9% lower odds of having more saved compared to those without such an account. This contradicts the findings of Heckman and Hanna (2015), who emphasized the positive influence of retirement savings on the likelihood of saving. One might expect that individuals with retirement savings accounts would have more savings, or savings closer to \$100k. This discrepancy presents a potential area for further research to understand why this relationship is different for lower-saving individuals.

4. Stocks, Bonds, ETFs, or Mutual Funds Outside Retirement Accounts.

Our results show that individuals with these assets have 49.1% lower odds of having less than \$100k saved compared to those without such resources, indicating that these assets are generally associated with higher savings levels. However, among individuals with less than \$100k in savings, those with these assets have 6.0% lower odds of having savings closer to \$100k compared to those without. This suggests that, although having these assets may be associated with higher overall savings, they do not seem to significantly increase savings within the group of individuals with lower savings. While existing literature

does not explore this predictor, it presents a potential area for further research due to its unexpected findings in the context of Americans with lower relative savings.

5. Emergency Expenses and Credit Cards.

Across savings groups, individuals who use credit cards for emergencies and pay them off over time have 1.66 times the odds of having less than \$100k in savings. However, when focusing only on those with less than \$100k saved, the results indicate that individuals who put an emergency expense on their credit card and pay it off over time have 1.05 times the odds of having savings closer to \$100k. These findings provide another area for further research as to why this relationship differs when studying savings behavior across versus within groups.

6. Education Level.

Comparing savings groups, the odds of having less than \$100k saved for individuals with a master's degree or professional degree are 44.0% and 56.0% lower, respectively, compared to those with less than a high school degree. These findings align with Yuh and Hanna (2010), which suggest that higher education is associated with better savings behavior. However, when focusing only on those with less than \$100k saved, individuals with some college or an associate's degree have 4.5% lower odds of having more saved compared to those without a high school degree. This finding contrasts with Yuh and Hanna (2010) and points to an area for further research to better understand why this relationship does not hold for those with lower savings.

7. Self-Employment.

According to the logistic regression results, the odds of having less than \$100k in savings for self-employed individuals are 36.6% lower than those for individuals working in government. This finding is consistent with Ramanathan (1969) and Yuh and Hanna (2010), who found that self-employed individuals tend to save more. However, the beta regression reveals that the odds of having savings closer to \$100k for self-employed individuals are statistically the same as for government workers. This contrasts with the findings of Ramanathan (1969). The differing relationship between self-employment and savings behavior across the two regressions presents another potential area for further research.

Below are the variables significant to only one model, or insignificant in both, but mentioned in some capacity in the above literature. We will also analyze these and their relation to previous studies.

1. Age.

When comparing savings groups, the odds of having less than \$100k in savings decrease by 1.66% with each additional year of age. This finding is consistent with Joseph (2019), who identified age as a significant factor in American savings behavior. However, age does not prove to be a significant predictor of savings behavior for individuals with less than \$100k saved, which aligns with the findings of Hogarth and Anguelov (2003).

2. Student Loan Debt.

Across groups, individuals with current student loan debt have 2.24 times the odds of having less than \$100,000 in savings. However, within the group of individuals with less than \$100k saved, this variable is not significant. While the literature referenced above does not discuss student loan debt, our findings offer a new perspective on understanding savings behavior across distinct groups.

3. Other Variables (Gender, Marital Status, Children, and Race/Ethnicity).

Variables such as gender, marital status, presence of children, and race/ethnicity were not significant predictors in our logistic regression. These findings are consistent with Khandu et al. (2021) for gender, Hogarth and Anguelov (2003) for marital status, and Heckman and Hanna (2015) for

race/ethnicity. Gender, marital status, and presence of children are also not significant in our beta regression, which aligns with Hogarth and Anguelov (2003). However, race/ethnicity, specifically the ‘Other’ ethnicity category, showed significance, as individuals in this group have 6.0% lower odds of saving closer to \$100k compared to White people. This finding contrasts with Heckman and Hanna (2015), but supports Joseph (2019) as there is evidence of racial/ethnic differences in savings behavior amongst Americans with lower (relative) savings amounts.

5. CONCLUSION

This paper examines the factors influencing American savings behavior in the 21st century. We conducted logistic and beta regressions to identify significant predictors of savings behavior across different savings categories (less than \$100k vs. more than \$100k) and within the less-than-\$100k category. In both cases, we excluded non-significant variables, resulting in the most parsimonious models. These methods enabled us to explore savings behavior from both categorical and continuous perspectives, providing insights that either aligned with, contrasted, or expanded on existing literature. The analysis also highlighted relationships between predictors and savings behavior that were consistent across both regressions, as well as those that differed.

We found overlap in the factors that are significant in understanding American savings behavior, both within and between groups. Income level, homeownership, retirement savings accounts, stocks, bonds, ETFs, or mutual funds outside retirement accounts, education level, and self-employment were significant predictors in both models. However, the impact of most variables varied when comparing across savings groups versus within the less-than-\$100k group. Only the income variable showed relatively consistent results across both the logistic and beta regressions. The other six shared predictors showed different effects on savings. These varying effects provide areas for further research aimed at uncovering why these relationships differ when comparing across versus within groups. Additionally, several predictors not addressed by existing studies offer new perspectives for understanding American savings behavior and provide a more nuanced view of the key factors influencing savings patterns.

Some variables are significant in only one of the models. Age and student loan debt were significant in logistic regression but not beta regression. These findings introduce new perspectives on savings behavior, as existing literature did not extensively discuss the impact of student loan debt. Its relationship with savings behavior could offer valuable insight into how debt impacts savings across different income groups and potential inhibitors of saving. Additionally, these results raise further paths of inquiry as to why these variables were not significant when focusing solely on Americans with less than \$100k saved.

Three variables—gender, marital status, and the presence of children—were not significant in either regression, despite being significant in at least one of the previous studies. These findings are noteworthy, as they challenge prior assumptions about the variables that influence savings behavior. It prompts areas of further investigation to understand why these variables are not significant in understanding American savings behavior in the 21st century.

This study offers new insights into the factors influencing American savings behavior in the 21st century, revealing both alignments with existing literature and the emergence of new variables that merit further examination. By considering both categorical and continuous models, we provide a more nuanced understanding of how numerous factors shape savings behavior in contemporary Americans.

References

- Callen, T., Thimann, C., IMF, & Asia. (1997). *Empirical Determinants of Household Saving: Evidence from OECD Countries* /: prepared by Tim Callen and Christian Thimann. United Nations Digital Library System; International Monetary Fund, <https://digitallibrary.un.org/record/256793?ln=en>
- Davis, E. P., & Schumm, W. R. (1987). *Savings Behavior and Satisfaction with Savings: A Comparison of Low- And High-Income Groups*. *Home Economics Research Journal*, 15(4), 247–256. <https://doi.org/10.1177/1077727x8701500405>
- Furnham, A. (1985). *Why Do People Save? Attitudes to, and Habits of Saving Money in Britain*. *Journal of Applied Social Psychology*, 15(5), 354–373. <https://doi.org/10.1111/j.1559-1816.1985.tb00912.x>
- Heckman, S. J., & Hanna, S. D. (2015). *Individual and Institutional Factors Related to Low-Income Household Saving Behavior*. *Journal of Financial Counseling and Planning*, 26(2), 187–199. <https://doi.org/10.1891/1052-3073.26.2.187>
- Hefferan, C. (1982). *Determinants and Patterns of Family Saving*. *Home Economics Research Journal*, 11(1), 47–55. <https://doi.org/10.1177/1077727x8201100109>
- Hogarth, J. M., & Anguelov, C. E. (2003). *Can the Poor Save?* *Journal of Financial Counseling and Planning*, 14(1). https://www.researchgate.net/publication/252256832_Can_the_Poor_Save
- Joseph, M. (2019). *Digital Commons at Buffalo State Economic and Demographic Determinants of the Consumption, Saving, and Borrowing Behaviors of Households*. https://digitalcommons.buffalostate.edu/cgi/viewcontent.cgi?article=1041&context=economics_theses
- Khandu, P., Dema, P., Dema, S., Wangmo, R., & Choden, Y. (2021). *International Journal of Innovative Science and Research Technology*, 6. [https://www.ijisrt.com/assets/upload/files/IJISRT21DEC207_\(7\)_\(1\)_\(1\).pdf](https://www.ijisrt.com/assets/upload/files/IJISRT21DEC207_(7)_(1)_(1).pdf)
- New National Poll: Economic Hardships of Millions of Middle-Class Americans Go Unseen, Most Believe They Will Continue to Struggle Throughout Their Lives*. (2024, June 4). *Cssny.org*. <https://www.cssny.org/news/entry/national-poll-economic-hardships-american-middle-class-true-cost-of-living-press-release>
- Ramanathan, R. (1969). *An Econometric Exploration of Indian Saving Behavior*. *Journal of the American Statistical Association*, 64(325), 90–101. <https://doi.org/10.2307/2283721>
- Shin, S. H., & Kim, K. T. (2018). *Perceived Income Changes, Saving Motives, and Household Savings*. *Journal of Financial Counseling and Planning*, 29(2), 396–409. <https://doi.org/10.1891/1052-3073.29.2.396>
- Yuh, Y., & Hanna, S. D. (2010). *Which Households Think They Save?* *The Journal of Consumer Affairs*, 44(1), 70–97. <http://www.jstor.org/stable/23859776>

The Pitcher's Dilemma

An analytical approach to identifying acute role differences

Nicholas Patrichuk – Hofstra University Hempstead NY 11549 patrichukn@gmail.com
Alexander Pelaez – Hofstra University Hempstead NY 11549 Alexander.Pelaez@hofstra.edu

Introduction

The dilemma of where to best utilize perceived starting pitchers and perceived relief pitchers has been an ongoing topic for years. This paper focuses on using different pitching statistics to improve the overall quality of a team's pitching staff, (i.e. Starters, Relievers) to identify pitchers along the margins that can maximize their efficiency within the pitching staff. If a team can succeed in maximizing their pitching staff, they will be presented with a competitive advantage among the other teams in their division and the league as a whole. Traditional baseball statistics and metrics put emphasis on physical characteristics such as throwing hard, and recording swing and misses throughout the game. However, these techniques do not analyze the subtleties of the skills that are nuanced for different types of pitching roles. The continued growth of analytics has not been met with open arms by all but it has allowed teams to make more factual decisions and not rely completely on intuition. While analytics attempts to identify strengths and weaknesses of players, positions in the field, or even strategical maneuvers, much less work is conducted around the subtle differences of roles within a group of players such as starting pitchers and relief pitchers.

This paper uses a combination of K-Means clustering and ordinal regression in order to identify characteristics of marginal players within one group of pitchers that could be optimal amongst the other group of pitchers. Specifically, we examined influential variables that would help endorse a role switch among pitchers helping to quell the question of mediocrity and realizing their full potential. Professional baseball has numerous instances with pitchers performing poorly in their traditional role only to excel in a new role. These types of switches add value on the team level and thereby increase the probability of winning. The methodology by identifying individual observations such as pitchers that are marginal within one group that could be optimal within another group could have significant value beyond sports

analytics. The combination of statistical techniques such as clustering and ordinal regression can provide other domains with a rich method of identifying these marginal observations.

Literature Review

The Current State of Pitching

The debate of starter to reliever value has been an ongoing topic as the two groups are completely co-dependent on one another (Whiteside et al, 2016). The current state of baseball is further enforcing this codependence as starters are pitching fewer innings year by year (Verducci 2023). Thus, relievers have been used far more frequently and this has put increased pressure on finding a good pool of relief pitchers, paving the way for an increased reliance on science and kinesics to potentially “mold” pitchers into a formidable figure. Donahue et al (2020) focused on the anthropometrics to see whether or not the physical makeup of each set of pitchers plays a part in which role they should serve in. The results of Donahue et al (2020) training determined that both position groups are similar in their makeups and score relatively close in physical performance check points. The training for both starters and relievers should be the same on a physical standpoint with a conscience approach based on rest. Assuming that the physical characteristics are the same, the mental acumen could be different amongst the two pitching groups.

A second study focused on the significance of upper body injuries for both starters and relievers. After results were calculated, it was found that the starting pitchers had a higher chance of arm injury in comparison to the relief pitchers (Bullock et al 2022). Due to the similarity of the anthropometrics, the study provides for the possibility that players would be able to fulfill alternate roles of the same position. If it is possible to identify more successful role changes, it could not only improve the performance of the team but it also could lower the risk of injury.

Positional alignments in the field are also another important technique that plays a role in a pitcher’s success role notwithstanding. Gerlica et al (2020) conducted a study focused on the realm of

fielding, which could have immense benefits for both starters and relievers throughout the game. The study ran a three-group K-Means cluster for outfield positioning which allowed position players to maximize the number of outs recorded in each game. A probability method was used to measure the positioning based on the centroid of each cluster as well as the hang time of a ball (Gerlica et al 2020). The shift of an outfield could go a long way in helping to preserving statistics like *FIP* (Fielding Independent Pitching), *BABIP* (Batting Average on Balls in Play), *ERA* (Earned Run Average), *WAR* (Wins Above Replacement) among others that are implemented in this analysis. Pitching stress factors like attending to runners leading off bases or trying to steal can be decreased as a result of the optimal alignments. While batted balls hit out of these alignments could prove to be problematic, the potential to save additional runs outweighs this risk. These benefits are further enhanced due to the dynamic nature of baseball and the increased adaptation of shifts.

Use of K-Means Clustering

K-Means clustering provides a unique perspective to identify different clusters within different groups, some of which may have some similarity. K-Means clustering has broad applicability across a number of domains including Finance and Healthcare. Momeni (2015) used K means clustering of financial ratios to accurately classify different industries. The paper attempts to classify the cement, metal and automotive parts industry based on stock prices, which could provide better opportunities for stakeholders. The introduction of large clusters helps paint a clearer picture for investors to truly understand the markets they target.

In the same manner, identifying pitchers in clusters is similar since there is a large pool of pitchers, however, only an acute amount of those pitchers could feasibly become optimal in the other pitching group. Just as Momeni (2015) was able to correctly segment these three industries, our pitching clusters will give an accurate depiction of the best pitchers to invest in. The clusters will show what pitching qualities, like financial ratios in Momeni (2015), will help equip the managers and owners, like

stakeholders in Momeni (2015), with the best possible pool of pitching to select from.

In addition, the use of K-Means clustering has been effectively used in the Healthcare industry with a focus on imaging. Using this segmentation through clustering, 3-D visualizations are enhanced to help detect any diseases that the subject may be at risk for. H.P Ng et al (2006) used K-Means clustering to create a foundation of the imaging before a second algorithm, the watershed method, was reintroduced with this segmentation. The specific images used in this study were 50 2-D T1 MR images of the head with three different parameter groups of thickness (H.P Ng et al 2006). These images identified that the K-Means clustering was a beneficial first step in analyzing these scans due to it lowering over-segmentation as a whole. The benefits of less over-segmentation are immense because the images produced are a closer representation of the body part they are examining. While the study mentions that over-segmentation decreased as a whole, it did not say K-Means clustering fully eliminated this concept. This demonstrates one of the shortcomings in the method due to the fact that randomness plays a part in creating the clusters. K-Means clustering may serve a researcher well initially; however, it should be used as a means to an end.

The performance of starters and relievers can be clustered in a likewise manner. Statistics withstanding, teams essentially rank pitchers in both the starting rotation as well as the bullpen. The starting rotations usually contain five-six pitchers depending on the preference of the club. The “ace”, #1 starter, is often the best pitcher of the group. They are the pitchers who will be called upon for big moments and playoff games. From there, the rest of the rotation is filled out generally in descending skill level with exceptions made for specific matchups based on the opposition. In the bullpen, the “closer” tends to be regarded as the best reliever as they are tasked with going out and getting crucial outs or converting saves. They are perceived to be equipped with the best pitching arsenal and are able to get hitters out at the highest rate no matter whether the bases are loaded or empty. Before the closer, the presence of a good setup pitcher holds immense value to the bullpen. The cluster “below” closers and setup relievers would be middle relievers and are generally the largest cluster in a reliever’s subset. The

last two clusters in a bullpen setting would include long relievers as well as injury replacements.

The underlying concern is whether the team is truly maximizing the value of the pitcher in their current role. The two subsets of starters and relievers may have marginal clusters in which a player in one of the marginal clusters could be of optimal value in another. The use of the clustering based on a specific characteristic allows one to determine where the marginal differences may be as opposed to advising the pitcher to switch their mechanics. Introducing a regression type model to this dilemma is also important because it is not practical to create the clusters and then try to label them based on comparing their general numbers to the mean of the cluster.

Use of Ordinal Regression

An additional method employed in this paper is ordinal regression and its ability to predict where certain pitchers land after a cluster analysis is performed in either group. Liu and Koirala (2012) used ordinal regression in an effort to predict mathematic proficiency levels. The authors used ordinal regression against other statistical methods to determine the viability of attaining better results. Education is one of the pillars of global society and the fact that ordinal regression was used in Liu and Koirala's research shows the immense benefit it could have to additional studies such as the Pitcher's Dilemma.

Ordinal regression has also been employed in determining the quality of life in patients. Abreu et al (2008) conducted a study centered around questionnaires that focused on certain components of a schizophrenic patient's life. The specific categories enclosed were centered on the domains of (1) social, (2) occupational and (3) interpersonal relations (Abreu et al 2008). Once questionnaire responses were recorded, they were ranked based on an ordinal scale. If the patient had a cumulative high score based on these questions, it signified a higher quality of life. After compiling the results of the questionnaires, the different ordinal regression methods were used to examine the data. Their results concluded that the ordinal regression models were useful in helping to analyze the quality of life for patients as long as sample size was specified. Similarly, using ordinal regression to predict the performance of pitchers, specifically across groups, could provide significant insight into the changing role of a pitcher.

Have This Process Been Repeated?

While the previous sections touch on the anatomy of pitchers as well as the methods used in the research, the next step was to find if this process was replicated from top to bottom in other areas. From a surface level, the goal in this research is to take two average or below average groups and switch their roles to increase both trajectory and value. When looking at other research, there were attempts to do something similar in another sect of the healthcare field, specifically motor learning. The study included 18 volunteers (10 males and 8 females) who were described as healthy and basically were similar from a medical standpoint (Balsley et al 2002). After the tracing results were completed, the numbers of clusters were determined based on the generalization error method which focuses on maximizing relationships based on predictions. Overall, the clusters showed an increase in visual and sensorimotor areas, which was attributed to the increase of hand and eye movement from the practice trials (Balsley et al 2002).

The similarity between these two pieces of research is based upon the type of repetitive methodology running against an uncertain outcome. Balsley et al (2002) focused on the drawing and traces of star-shaped images over a set period of time. The action of pitching focuses on the use of both the upper and lower body to throw a baseball over the plate and keep the batter from putting the ball in play. The motor research conducted by Balsley et al (2002) used a centering system in their clustering to populate where the functions were most prevalent during each trial. In the baseball research, the “cluster center” would be the optimal cluster for both the starters and the relievers, where the characteristics parsed were maximized and could be used to help make an argument for a pitcher to switch roles. The “special extent” of the cluster in the pitcher’s data would just be based on the number of subjects in each group. The goal when clustering pitchers was to have as close to a homogenous number of subjects in each original cluster. The ordinal regression produced varying magnitudes of size which was deemed ok because this was ultimately the result of the cluster. It would not have been beneficial if the clusters remained synonymous post regression because the results would not have been as telling.

Methodology

The dataset used in this analysis was the Fangraphs pitching data from the prior three seasons (2021, 2022, 2023). The total observations in this dataset were 3,325. The first task was to break this data down into a smaller subset that would truly allow for relationships among the position groups to be discovered. In the “Starters” subset, pitchers who qualified for the ERA statistic were used (160 innings). These specifications are made for starters because it eliminates outlier pitchers who may have been used for a small number of innings and have a 1.00 or lower ERA. In the “Relievers” subset, all pitchers having pitched at least 55 innings were included. The rationale for picking 55 innings was based on the Trevor Hoffmann and Mariano Rivera Reliever of the Year awards. These are the awards that honor the perceived best reliever in both the American and National League. Past winners of these awards have not pitched less than that number which indicates the pitcher was able to stay healthy while handling ample usage. The condensing of the subsets also eliminated the presence of outlier candidates who pitched only a handful of times. Whether it is a player who gets called up from the minor leagues or gets injured, the goal was to avoid appearances that may skew the data in either direction. As previously mentioned with the starter’s distinction, relievers could have statistic fluctuation and it is imperative to balance this variance.

The next step in the process was to begin to cluster the data based on a condensed group of variables. In the Fangraphs data, there were 302 variables encompassing the entire dataset with some variables including little to no values. Condensing the variables in the Pitchers data allowed us to analyze metrics that would be applicable across both groups. This led to statistics such as “Saves” for relievers and “Wins” for starters being removed because the probability of recording this stat was much higher for the incumbent group as opposed to the inverse group.

Once the repetitive and unnecessary variables for this research were diluted, the next objective was to identify the variables for which would be the most beneficial for a cluster analysis. It was important to highlight both their tendencies as well as the way in which they handled high leverage

situations. Variables such as *Strikeout Rate* and *Walk Rate* were included as well as contact outcomes such as *Groundball Percentage*, *Flyout Percentage* and *Pop Out Percentage*. The aforementioned values gave a full display of each specific pitcher's tendencies, both strong and weak. Other higher leverage statistics also included were: *Left on Base Percentage*, *Exit Velocity*, *Hard Hit Rate*, *FIP* (Fielding Independent Pitching) and *Launch Angle*. These are all properties that demonstrate the specific way in which the batters are making contact on the pitches that are thrown within each at bat. Other variables were included such as: *Hits per 9 Plus*, *Average Plus*, *Strikeout to Walk Rate Plus* and *WAR* (Wins Above Replacement) needed to be altered in order for them to function in the cluster analysis. Box plots were used as a visual to see if the dispersion of starters and relievers had more outliers based on the scale key used. From a statistical standpoint, Wilcoxon Tests were used as they compare the location of two different populations based on the samples of the data. Since all of the aforementioned variables are on a different scale, they were normalized using a Z-score.

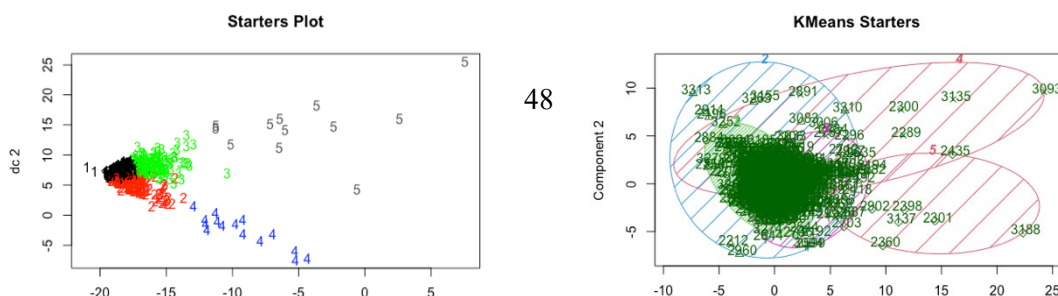
After solidifying the clusters, the next step was to create an accurate model that would predict where each position group would fall out of the five clusters for the opposite group. If the opposite role fell into the optimal cluster, they would be a candidate to change roles. Each of the models were determined by the variables that were the most significant in changing the output produced by the system. There were multiple candidates who spawned in the optimal clusters across both position groups. In order to select the players to highlight, some common knowledge including injury history, innings pitched and age were factored in to support the viability of a role change.

Results

The models to predict where a specific pitcher would land in the opposite cluster are listed as follows:

Model to predict Starters based on Relievers clusters: $GB\_pct\_plus + WHIP\_plus + WAR\_Z$

Model to predict Relievers based on Starters clusters: $Pull\_pct\_plus + FIP + Oppo\_pct\_plus$



These were selected as the respective variables due to the fact that they are the significant variables found in each truncated dataset. Both of the models contain three variables each as the addition of extra created multicollinearity effects as well as issues with processing. There is a case to be made for each grouping that will be mentioned later within this section.

When looking at the first model for predicting where relievers landed in starter clusters, the variables selected were *GB_pct_plus*, *WHIP_plus* and *WAR_Z*. The first metric essentially is an aggregated way of describing how often a pitcher gets a hitter to ground out into the field of play. The groundball rate could be an important factor in predicting Reliever-to-Starter viability because of the two unique molds that are shaping with relievers. These molds include being a strikeout pitcher or a “specialist” who can be brought in to face hitters with specific handedness or tendencies. The second metric found in this model was *WHIP_plus* which is summation of hits and walks divided by innings pitched. *WHIP_plus* is an important measure to predict this role switch because their sample size will go up naturally in a starter’s role. If the pitcher in question has a great *WHIP_plus*, chances are it could signal more success when the sample size opportunities increase. The last metric found in this model was *WAR_Z*, which we created to normalize the original metric, *WAR*. *WAR* accounts for how much better a specific player may be from the general replacement or “average” player. *WAR* could be viewed as important for this research as starters on average have a higher *WAR* than relievers. If a reliever has a substantial *WAR*, they could be in a prime position to succeed as a starter.

On the other hand, the second model for predicting where starters landed in the reliever clusters included: *Pull_pct_plus*, *FIP* and *Oppo_pct_plus*. *Pull_pct_plus* is a metric that focuses on the frequency that the batters are hitting pitches to the pull side. The pull side would be the left side for right-handed hitters and right side for left-handed hitters. This metric could be viewed as significant due to the heightened importance of outcomes in a relief role. Additionally, dimensions of the ballpark could play a part in this as the pull side could be larger than the opposite side. A fly ball to the pull side could result in

a run scoring via a sacrifice fly that may not happen if hit were hit to the opposite side. The second metric included in this model was *FIP* or Fielding Independent Pitching. This value focuses on the pitcher's true value and leaves out any errors or plays that are not the direct responsibility of the pitcher. The significance lies in the potential for increased defense in the later innings. Earlier in games, teams are focused on scoring runs and once a lead is secured, there is more of an emphasis on upholding it. The last metric included was *Oppo_pct_plus* which is essentially the antonym of *Pull_pct_plus*. The opposite field is right field for right-handed hitters and left field for left-handed hitters. While both variables are different, the aforementioned rationale for *Pull_pct_plus* can apply for both.

One reliever who populated in the optimal starter cluster based on the model was Nick Sandlin, a reliever who has spent his whole current career with the Cleveland Guardians. As a reliever, he has never started a game and has accumulated 184 innings over his four-year career. The season in which he qualified for the 55-inning threshold came in 2023 where he pitched 60 innings. Sandlin was a higher draft pick going in the second round of the 2018 draft and reached the Major Leagues after three years.

The Guardians have opted to use Sandlin as a reliever throughout his entire career in the organization with the exception of starting one game this year in the way of an "opener". An "opener" is described as a reliever who starts the game but only pitches one to two innings. After doing the first cluster of analysis amongst just relievers, Sandlin populated as a Cluster 3 reliever. As mentioned at the onset of this research, Cluster 3 relievers were viewed as "middle relievers" or hovering around average. Once the model was created to predict where relievers would land amongst the starter cluster, Sandlin populated in Cluster 1. This demonstrates that the model provides evidence that Sandlin could present more value to his respective team as a starter as opposed to a reliever. After some additional background research on Sandlin, it was discovered that he once was a starter during his collegiate career. During his final season at Southern Mississippi University, Sandlin was a starter and pitched remarkably to the tune of a 1.06 *ERA* in 101 innings. Whether it is with the Guardians or another team, experimenting with Sandlin as starter could be an advantageous endeavor. There are other potential candidates who could fall

into this category and will allow teams to both save money in their budget while also potentially increasing ticket size with more success.

Conclusion

The preliminary analysis regarding this concept showed that the use of cluster analysis paired with ordinal regression can be used to identify acute differences in one position role and endorse their switch to the other role. While there were several pitchers who moved to a different cluster, the primary focus was on finding those who offered the most significant change and value. Nick Sandlin was identified as an example candidate who initially started in Cluster 3 of his original group and then populated in Cluster 1 of the other group through the use of our model.

Additional papers could include other pitchers who fell into these confines where a role change could be supported. There are several other candidates but it is difficult to do an analysis on each one in a singular paper. In addition, the methodology and outside research demonstrated that this type of clustering into modeling could be used in other industries as well. With this research, we hope to have exemplified to both companies and teams alike that acute differences among members of their respective groups or teams could be identified while providing a supportive layout for change.

References

- Abreu, M. N. S., Siqueira, A. L., Cardoso, C. S., & Caiaffa, W. T. (2008). Ordinal logistic regression models: application in quality of life studies. *Cadernos de Saúde Pública*, 24, s581-s591.
- Balslev, D., Nielsen, F. Å., Frutiger, S. A., Sidtis, J. J., Christiansen, T. B., Svarer, C., ... & Law, I. (2002). Cluster analysis of activity-time series in motor learning. *Human brain mapping*, 15(3), 135-145.
- Bullock, G. S., Thigpen, C. A., Collins, G. S., Arden, N. K., Noonan, T. K., Kissenberth, M. J., & Shanley, E. (2022). Hazard of arm injury in professional starting and relief pitchers. *Journal of Athletic Training*, 57(1), 65-71.

- Donahue, P. T., Beiser, E., Wilson, S. J., & Garner, J. C. (2020). Comparison of anthropometrics and physical performance in professional baseball pitchers. *Journal of Trainology*, 9(2), 39-42.
- Gerlica, J., LaDuke, I., O'Shea, G., Pluemer, P., & Dulin, J. (2020). Quantifying the Outfield Shift Using K-Means Clustering.
- Liu, X., & Koirala, H. Ordinal Regression Analysis (2012): Using Generalized Ordinal Logistic Regression Models to Estimate Educational Data. *Shlomo S. Sawilowsky*, 242.
- Momeni, M., Mohseni, M., & Soofi, M. (2015). Clustering stock market companies via k-means algorithm. *Arabian Journal of Business and Management Review (Kuwait Chapter)*, 4(5), 1-10.
- Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2006, March). Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE southwest symposium on image analysis and interpretation* (pp. 61-65). IEEE.
- Verducci, T. (2023, May 22). *Starting pitching's downward spiral shows no sign of slowing*. SI. <https://www.si.com/mlb/2023/05/22/starting-pitching-decline-pitch-count-timer-injuries>
- Whiteside, D., Martini, D. N., Zernicke, R. F., & Goulet, G. C. (2016). Changes in a Starting Pitcher's Performance Characteristics Across the Duration of a Major League Baseball Game. *International journal of sports physiology and performance*, 11(2), 247–254. <https://doi.org/10.1123/ijsp.2015-0121>

A MACHINE LEARNING ALGORITHM BASED ON MAX-NORMALIZED EFFICIENCY-ORIENTED EVALUATION METHOD FOR MEASURING PERFORMANCE

Jae-Dong Hong, Industrial Engineering, South Carolina State University,
Orangeburg, SC 29117, 803-536-8861, jhong@scsu.edu

ABSTRACT

This paper proposes a machine learning algorithm based on an efficiency-oriented evaluation model with a max-normalizing approach to evaluate the performance of decision-making units (DMUs) in the context of data envelopment analysis (DEA). The three critical barriers for DEA to become a tool for Big Data-enabled analytics are excessive computational times for a large set of DMUs, poor discriminatory power, and illogical and inconsistent ranking results. Through three numerical examples, we demonstrate that the proposed approach can eliminate such barriers and allow the decision-makers to evaluate and rank DMUs more consistently and faster than DEA-based models.

Keywords: Machine learning algorithm, Efficiency-oriented evaluation, Max normalizing, Data envelopment analysis, Decision-making unit

INTRODUCTION

Big Data research has become an essential area of operations analytics. Various researchers show that Big Data can be applied to improve company productivity or efficiency and will be vital for enterprises to grow and achieve a competitive advantage. Charles et al. (2021) provide an overview of the current avenues of research for the studies aimed at integrating data envelopment analysis (DEA) with Big Data. Zhu (2022) also suggests that DEA should be viewed as a tool or method for data-oriented analytics in performance assessment and benchmarking. Castellano et al. (2023) propose an innovative technique to enhance the accuracy of productivity measures through the use of Big Data in a cluster analysis combined with DEA. When researchers attempt to integrate DEA into Big Data, the first serious barrier between Big Data and DEA is the excessive computational time for solving a large set of DMUs in the Big Data context, as Barr and Durchholz (1997) and Khezrimotlagh et al. (2019) show in their papers. Muren et al. (2024) propose a parallel framework algorithm for solving large-scale DEA models, claiming their algorithm has substantial advantages in decreasing computational time. They show that their method reduces the computational time by at least 12.52%, comparing the framework method that Khezrimotlagh et al. (2019) propose.

Among performance evaluation and benchmarking methods, data envelopment analysis (DEA) has emerged as the most popular technique that uses Linear Programming (LP) to compute the relative performance of a set of peer decision-making units (DMUs) by comparing how well the DMU uses its inputs to produce outputs. The conventional/traditional DEA (C-DEA), introduced by Charnes et al. (1978), eventually determines which DMUs attain the most efficient outcome using given inputs and which do not. The C-DEA model generates a single, comprehensive

performance measure, an efficiency score (ES), for each DMU. The C-DEA model classifies all DMUs into two groups, which would separate efficient units from inefficient units. The C-DEA frequently produces too many efficient DMUs out of all DMUs under evaluation, showing poor discriminatory power, so there is no way to rank efficient DMUs. However, identifying the best DMUs is frequently more important than maximizing the individual ES in many practices.

To remedy this deficiency of C-DEA, Sexton et al. (1986) suggest the cross-evaluation DEA (CE-DEA) method to evaluate and rank DMUs, with the main idea of using C-DEA to do the peer evaluation rather than C-DEA's pure self-evaluation. Due to its enhanced discriminatory power, the CE evaluation has generated a significant number of applications in the DEA literature (see Liang et al., 2008; Wang & Chin, 2010; Gavgani & Zohrehbandian, 2014; Hou et al., 2018; Lee, 2019; Liu et al., 2019; Hong & Jeong, 2020; Hong, 2022; Ganji et al., 2023; Ravanos & Karagiannis, 2024). As Doyle and Green (1994) indicated, the non-uniqueness of CE scores (CESs) often results from alternative optimal weights in the C-DEA model, implying that the CESs depend on the optimization software used. The rankings generated by CE-DEA are neither robust nor consistent. The idea of super-efficiency (SE), developed by Anderson and Peterson (1993), is that the C-DEA model is applied, excluding a DMU under evaluation from the reference set of the C-DEA model. Nayeibi and Lotfi (2016) describe ranking methods based on SE. The critical issue of using the SE model is that the two adjacent DMUs decide the SES of an efficient DMU, so it would sometimes be awkward for some DMUs to be ranked by the SES. As a result, the rankings generated by S-DEA frequently turn out to be illogical or unreasonable.

Seiford and Zhu (1999) propose the stratification/context-dependent DEA (S-DEA) method to classify all DMUs into several efficiency levels. The S-DEA would measure the attractiveness (progress) score of DMUs at a higher (lower) level when DMUs at a lower (higher) level are chosen as the evaluation background/context. These attractiveness/progress scores are inconsistent as the evaluation context changes due to the DEA model's intrinsic deficiency.

Machine learning (ML) means instructing a machine or an algorithm to perform tasks. It is a subfield of artificial intelligence (AI), an automated process extracting patterns from data. Malone et al. (2020) state, "The function of an ML system (MLS) can be descriptive, implying that the MLS uses the data to explain what happened; predictive, meaning the MLS uses the data to predict what will happen; or prescriptive, implying the MLS will apply the data to make suggestions about what action to take." Kelleher et al. (2015) state, "Machine learning algorithm (MLA) automates the process of learning a model that captures the relationship between the descriptive features and the target feature in a dataset. MLA works by searching through a set of possible prediction models for the model that best captures the relationship between the descriptive features and target features in a dataset. When using ML, analysts begin with a model to work with, supply the data, and let the computer model train itself to find patterns and/or make predictions. Over time, the analyst may tweak the model to improve its performance (Kultur & Calayan, 2017).

This paper proposes MLA based on the efficiency-oriented evaluation (ML-EOE) model with the max normalizing (MXN) approach. The proposed method would allow the system to learn about automation quickly. When new DMUs are added to the existing DMU set, all DEA-based

methods require reevaluating all DMUs, including newly introduced DMUs. In contrast, the proposed method allows the system to assess the new DMU only if the maximum input and output data do not exceed the previous maximum values. Not like DEA-based methods requiring optimization software, the proposed method does not require any software but simple computations. Thus, computational times would be almost negligible compared to those needed by DEA-based methods.

Decision-makers are usually interested in selecting the top-ranked DMU(s) before choosing the efficient ones among DMUs under evaluation. If the #1 ranked DMUs depended on which DEA method is applied, it would confuse the decision-makers or practitioners. The proposed approaches would be better than the DEA-based models for answering the age-old question, ‘Which DEA ranking method should we use?’ since it can eliminate the DEA’s intrinsic critical weaknesses from the self-evaluation principle, including long computational times for a large set of DMUs in the Big-Data context.

DEA-BASED EVALUATION METHOD

The ratio-form or fractional DEA model uses the ratio of weighted outputs to weighted inputs to estimate the relative efficiency of DMU_j as an objective function to be assessed relative to the ratios of all DMUs. The fractional DEA model is stated (Zhu, 2014):

Objective Function: Maximize the efficiency rating θ for DMU_j

$$\max \theta_j = \frac{\sum_{r=1}^s u_{rj} y_{rj}}{\sum_{i=1}^m v_{ij} x_{ij}}, \quad j = 1, 2, \dots, n, \quad (1)$$

which is subject to the constraint that when the same set of multipliers, u and v , is applied to all other DMUs to be compared, all DMUs will be less than or equal to 100% efficient as follows:

$$\frac{\sum_{r=1}^s u_{rj} y_{rw}}{\sum_{i=1}^m v_{ij} x_{iw}} \leq 1, \quad \forall j \text{ and } w = 1, 2, \dots, n, \quad (2)$$

$$u_{rj}, v_{ij} \geq 0, \quad r = 1, \dots, s; \quad i = 1, \dots, m; \quad w = 1, 2, \dots, n,$$

where

$j = DMU_j$ being evaluated in the DEA analysis, $j = 1, \dots, n$

y_{rj} = amount of output r generated by DMU_j

x_{ij} = amount of input i used by DMU_j

s = number of outputs produced by DMUs

m = number of inputs used by DMUs

u_{rj} = multipliers or weight assigned by DEA to output r for DMU_j

v_{ij} = multipliers or weight assigned by DEA to input i for DMU_j

Based on the fractional DEA model in (1)-(2), the conventional/traditional DEA (C-DEA) and cross-efficiency DEA (CE-DEA) models were developed (see Cooper et al., 2007). The CE-DEA method, which consists of two phases, was proposed to rank DMUs with the central idea of using DEA to do peer assessment rather than pure self-assessment (see Sexton et al., 1986). The

weights or multipliers from the first phase are used for all DMUs to get the cross-efficiency score (CES) for each DMU in the second phase. In the first phase, the ES of DMU_j is obtained. In the second phase, the other DMUs are evaluated using the optimal weights of a rating DMU_j .

The SE-DEA would produce a super-efficiency score (SES) obtained from the C-DEA model after a DMU under evaluation is excluded from the reference set (see Anderson and Peterson, 1993). In the SE method, the frontier line generated from the remaining DMUs changes for each efficient DMU to be evaluated, so the SESs of efficient DMUs can have higher values than one (1.000), the maximum value in ES obtained by other DEA methods.

Seiford and Zhu (1999) propose that S-DEA be used to measure the attractiveness and progress scores of DMUs in a given evaluation context. See Seiford and Zhu (1999) for a detailed procedure to stratify DMUs into different levels and compute the attractiveness and progress score of DMUs at each level, ℓ . The average score of attractiveness and progress (ASAP) is applied to rank DMUs.

THE EOE METHOD WITH MAX NORMALIZING (MXN-EOE)

If the magnitude of inputs/outputs is significantly different from each other, the EOE would not measure DMUs' performance correctly and effectively. For example, input #1 is the number of workers, such as several hundred workers, while input #2 denotes the total operating costs, such as several million dollars. Then, the magnitude of input #2 will dominate over that of input #1. Sortiros and Despotis (2012) introduce the max-normalized DEA by dividing the input/output by the maxima and showing that the max-normalized DEA model is structurally identical to the C-DEA model with un-normalized data. Such normalization will prevent round-off errors that imbalanced input/output may bring about. Following them (2012), this paper proposes the max normalizing (MXN) approach, where the input and output data are divided by each maxima. The MXN model also finds the same constant returns to scale (CRS) ES yielded by C-DEA. Now, the max normalized input i , x'_{ij} , and output r , y'_{rj} , for DMU_j are given by

$$x'_{ij} = \frac{x_{ij}}{\text{Max}_{\forall k} \{x_{ik}\}}, \forall i \text{ and } k = 1, 2, \dots, j, \dots, n \quad (3)$$

and

$$y'_{rj} = \frac{y_{rj}}{\text{Max}_{\forall k} \{y_{rk}\}}, \forall r \text{ and } k = 1, 2, \dots, j, \dots, n. \quad (4)$$

The efficiency score in (1) of DMU_j for the DEA method can be decomposed into

$$\text{Max } \theta_j = \frac{u_{1j}y_{1j}}{\sum_{i=1}^m v_{ij}x_{ij}} + \frac{u_{2j}y_{2j}}{\sum_{i=1}^m v_{ij}x_{ij}} + \dots + \frac{u_{sj}y_{sj}}{\sum_{i=1}^m v_{ij}x_{ij}}, \forall j. \quad (5)$$

Note that the poor discriminating power of DEA comes from the fact that unfavorable input/output data can be ignored by making the corresponding multipliers, u_{rj} and/or v_{ij} , become zeroes. In fact, many authors have pointed out mistreating input/output data as a critical weakness of the DEA models and have published various complementing methods. However, no

superior or universal method has appeared, as Aldamak and Zolfaghari (2017) observe. In contrast to the DEA methods, the proposed evaluation method treats all input/output fairly or equally regardless of their contribution to the ES. Instead of the multipliers, we use the proportion of each input/output to contribute to the evaluation. Let α_r and β_i , $\forall r$ and i , denote the proportion/weight of the contribution of output r and input i .

To find the average MXN input for DMU_j , we add the weighted inputs and then divide the number of inputs as follows:

$$I_j^{MX} = \frac{1}{m} \sum_i^m \beta_i x'_{ij}. \quad (6)$$

Similarly, the average MXN output for DMU_j , O_j^{MX} , is expressed as

$$O_j^{MX} = \frac{1}{s} \sum_r^s \alpha_r y'_{ij}. \quad (7)$$

Then, an unnormalized ES for DMU_j , ES_{jU}^{MX} , is written as the ratio of O_j^{MX} in (7) to I_j^{MX} in (6) as follows:

$$ES_{jU}^{MX} = \frac{\frac{1}{s} \sum_r^s \alpha_r y'_{ij}}{\frac{1}{m} \sum_i^m \beta_i x'_{ij}}. \quad (8)$$

Constraint (2) in C-DEA formulation ensures that an efficiency score (ES) does not exceed one (1.000). Similarly, we multiply ES_{jU}^{MX} in (8) by the inverse of the maximum value of ES_{kU}^{MX} to ensure that the normalized ES, ES_{jN}^{MX} , does not exceed one (1.000) as follows:

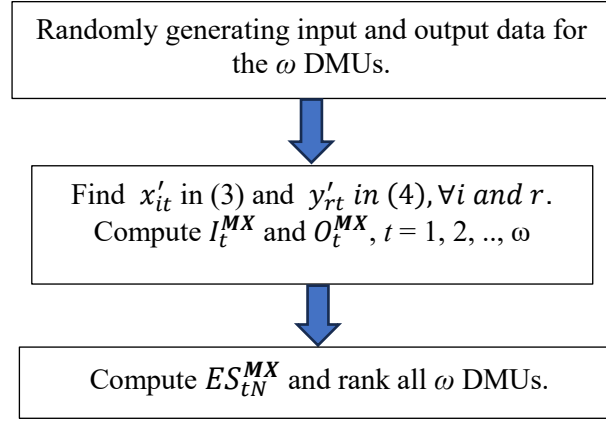
$$ES_{jN}^{MX} = \left\{ \frac{\frac{1}{s} \sum_r^s \alpha_r y'_{ij}}{\frac{1}{m} \sum_i^m \beta_i x'_{ij}} \right\} * \min_{\forall k} \left\{ \frac{\frac{1}{m} \sum_i^m \beta_i x'_{ik}}{\frac{1}{s} \sum_r^s \alpha_r y'_{ikj}} \right\}, k = 1, 2, \dots, n. \quad (9)$$

As described before, the CE scores generated by CE-DEA may not be unique due to multiple optimal weights/multipliers for inputs and outputs obtained from solving LP. It implies that the resulting CE scores depend on the applied software. In contrast, the efficiency score generated by the proposed method, ES_{jN}^{MX} , do not require any LP formulation on which all DEA methods are based. Hence, the proposed methods do not depend on the multiple optimal weights but on the output ratio to the sum of the inputs. The MXN-EOE is applied to eliminate the imbalance in the raw input/output data caused by units of measurement or different orders of magnitude. Now, we propose the following procedure of MLA based on the MXN-EOE:

The Procedure of Machine Learning Algorithm (MLA)

We select ω training DMUs with randomly generated inputs and outputs and put them in the training data. Figure 2 shows the steps of the training stage based on MXN-EOE as follows:

Figure 2: Flowchart of the training stage



{Training Stage}

Step 1: Generate input and output data for the η DMUs.

Step 2: Find x'_{it} and y'_{rt} , $\forall i$ and r from η DMUs. Compute I_t^{MX} and O_t^{MX} , $t = 1, 2, \dots, \eta$

Step 3: Compute ES_{tN}^{MX} and rank η DMUs based on ES_{tN}^{MX} .

{Initialization}

Step 1: Set $j = 1$. Set $x_i^M = x_{ij}$, $\forall i$, and $y_r^M = y_{rj}$, $\forall r$. Set $x'_{ij} = 1$, $\forall i$, and $y'_{rj} = 1$, $\forall r$. Set

$$ES_{jN}^{MX} = 1, \forall j, \text{ and } ES_{Max}^{MX} = 1.$$

Step 2: Set $j = j + 1$, $\lambda I = 0$, and $\lambda O = 0$. Enter DMU_j in the evaluation set, $k = \{1, 2, \dots, j\}$.

Step 3: If $x_{ij} \leq x_i^M$, set $x'_{ij} = \frac{x_{ij}}{x_i^M}$, $\forall i$. Otherwise, replace x_i^M with x_{ij} and set $x'_{ik} = \frac{x_{ik}}{x_i^M}$, $\forall k$, and

$$\lambda I = 1. \text{ If } y_{rj} \leq y_r^M, \text{ then } y'_{rj} = \frac{y_{rj}}{y_r^M}, \forall r. \text{ Otherwise, replace } y_r^M \text{ with } y_{rj} \text{ and set } y'_{rk} = \frac{y_{rk}}{y_r^M}, \forall k, \text{ and } \lambda O = 1.$$

{MXN-EOE Method}

Step 4: If $\lambda I \neq 0$ and $\lambda O \neq 0$, go to Step 5. Otherwise, compute ES_{jU}^{MX} in (19). If $ES_{jU}^{MX} \leq ES_{max}^{MX}$,

then compute $ES_{jN}^{MX} = \frac{ES_{jU}^{MX}}{ES_{max}^{MX}}$. Otherwise, replace ES_{max}^{MX} with ES_{jU}^{MX} and compute

$$ES_{kN}^{MX} = \frac{ES_{kU}^{MX}}{ES_{max}^{MX}}, \forall k. \text{ Go to Step 6.}$$

Step 5: If $\lambda I = 1$, compute I_k^{MX} in (17), $\forall k$. If $\lambda O = 1$, compute O_k^{MX} in (18), $\forall k$. Compute ES_{kU}^{MX} in (19), $\forall k$. Setting $ES_{max}^{MX} = \max_k \{ES_{kU}^{MX}\}$, compute $ES_{kN}^{MX} = \frac{ES_{kU}^{MX}}{ES_{max}^{MX}}$.

Step 6: Rank DMU_k based on ES_{kN}^{MX} , $\forall k$. Go to Step 2

To compare the results of the MXN-EOE methods with DEA-based models, we set α_r and β_i , $\forall r$ and i , equal to 1, since DEA-based models treat all inputs and outputs equally by default before they evaluate them. We use three well-known numerical examples in the DEA-related

literature to assess the performance of the proposed EOE methods and compare the results with those generated by the DEA-based models.

NUMERICAL EXAMPLES

Example 1. First, we consider the numerical example that Liang et al. (2008) illustrate with five DMUs. The data for each DMU consists of three inputs (x_{1j}, x_{2j}, x_{3j}) and two outputs (y_{1j}, y_{2j}) as shown in Table 1. To show the procedure of MLA, Table 2 shows the change of I_j^{MX} , O_j^{MX} , ES_{jU}^{MX} , and ES_{jN}^{MX} as a DMU has been added to the evaluation set.

Table 1: Five decision-making units

DMU	Input			Output	
	x_{1j}	x_{2j}	x_{3j}	y_{1j}	y_{2j}
1	7.0	7.0	7.0	4.0	4.0
2	5.0	9.0	7.0	7.0	7.0
3	4.0	6.0	5.0	5.0	7.0
4	5.0	9.0	8.0	6.0	2.0
5	6.0	8.0	5.0	3.0	6.0

We start with two DMUs, DMU_1 and DMU_2 . DMU_2 has a lower I_j^{MX} and a greater O_j^{MX} , so its ES_{jN}^{MX} is a perfect value of 1.000 with a #1 rank. Then, DMU_3 is added to the evaluation set, and its maximum input and output data values do not exceed the previous maximum values of DMU_1 and DMU_2 . Thus, I_j^{MX} , O_j^{MX} , and ES_{jU}^{MX} of DMU_1 and DMU_2 are not changed. The unnormalized ES of DMU_3 becomes the greatest DMU with $ES_{jU}^{MX} = 1.3170$. The normalized ES, ES_{jN}^{MX} , ranks $\{DMU_3, DMU_2, DMU_1\}$ as $\{#1, #2, #3\}$. Next, DMU_4 is added to the evaluation set, which has the greatest x_{3j} of 8.0, so adding DMU_4 to the evaluation set changes I_j^{MX} for all DMUs. Lastly, the data set of DMU_5 does not change any parameters.

Table 3 shows that DEA-based methods consistently rank the two efficient DMU_3 and DMU_2 as #1 and #2 (#1 by S-DEA), but CE- and S-DEA rank the three inefficient DMU_1 , DMU_4 , and DMU_5 differently from SE-DEA.

For further investigation, the three inefficient DMUs, after excluding two efficient DMUs, are evaluated and reported in Table 4. Since the two efficient DMUs are excluded, C-DEA rates these three DMUs as efficient with an ES of 1.000. Then, S-DEA classifies all three DMUs into one level, so ASAP is not applicable. Table 4 shows the ranks and the expected ranks, denoted by [R], based on the rankings in Table 3. Table 4 reveals that SE-DEA and MXN-EOE rank these DMUs consistently, but CE-DEA does not. Table 4 also shows that two DEA-based models, CE- and SE-DEA, rank DMU_4 higher than DMU_5 or DMU_1 , whereas the proposed method ranks DMU_5 higher than DMU_1 and DMU_4 .

Table 2: Change of efficiency scores and ranks by two through five DMUs

No.	DMU	I_j^{MX}	O_j^{MX}	ES_{jU}^{MX}	ES_{jN}^{MX}	R
1	1	0.9259	0.5714	0.6171	0.5583	2
	2	0.9047	1.0000	1.1052	1.0000	1
2	1	0.9259	0.5714	0.6171	0.4685	3
	2	0.9047	1.0000	1.1052	0.8391	2
	3	0.6507	0.8571	1.3170	1.0000	1
3	1	0.8843	0.5714	0.6462	0.4682	3
	2	0.8631	1.0000	1.1586	0.8394	2
	3	0.6210	0.8571	1.3802	1.0000	1
	4	0.9048	0.5714	0.6316	0.4576	4
4	1	0.8843	0.5714	0.6462	0.4682	4
	2	0.8631	1.0000	1.1586	0.8394	2
	3	0.6210	0.8571	1.3802	1.0000	1
	4	0.9048	0.5714	0.6316	0.4576	5
	5	0.7903	0.6429	0.8134	0.5893	3

Table 3: Comparison of diverse efficiency scores and rankings by five DMUs

DMU	DEA-Based Method								ML-EOE	
	C-DEA		CE-DEA		SE-DEA		S-DEA		MXN-EOE	
	ES	R	CES	R	SES	R	ASAP	R	ES_{jN}^{MX}	R
1	0.6857	5	0.5191	4	0.6857	5	0.642	4	0.4682	4
2	1.0000*	1	0.9161	2	1.1200	2	1.333	1	0.8394	2
3	1.0000*	1	0.9571	1	1.5000	1	1.333	1	1.0000	1
4	0.8571	3	0.6985	3	0.8571	3	0.857	3	0.4576	5
5	0.8571	3	0.4942	5	0.8571	3	0.600	5	0.5893	3

*: Efficient by C-DEA; R: Rank

Table 4: Comparison of efficiency scores and rankings for three inefficient DMUs

DMU	DEA-Based Method											ML-EOE		
	C-DEA		CE-DEA			SE-DEA			S-DEA			MXN-EOE		
	ES	R	CES	R	[R]	SES	R	[R]	ASAP	R	[R]	ES_{jN}^{MX}	R	[R]
1	1.000	1	0.771	3	2	1.123	3	3	N/A	N/A	2	0.794	2	2
4	1.000	1	1.000	1	1	2.100	1	1	N/A	N/A	1	0.774	3	3
5	1.000	1	0.805	2	3	2.100	1	1	N/A	N/A	3	1.000	1	1

R: Rank; [R]: Expected Rank

As DEA-based models implicitly assume equal weight to each input and output before evaluation, it is unreasonable for two DEA-based models to rank DMU_4 higher than DMU_5 . DMU_4 and DMU_5 have the same minimum input, 5.0, and the maximum output, 6.0. However, DMU_4 has the other two inputs, 9.0 and 8.0, which are greater than or equal to those of DMU_5 , 6.0, and 8.0. Except for the same maximum output value of 6.0, DMU_5 has a higher output value of 3.0 than 2.0 for DMU_4 , implying that DMU_5 should be ranked higher than DMU_4 . Thus, **the**

ranks generated by the proposed method would be more rational and logical than those by DEA-based models.

We implement the MXN-EOE method in an Excel spreadsheet with VBA (Visual Basic for Applications) on Intel® Xeon® Gold 5122 (2 processors) HP Z8 G4 Workstation PC with 32GB of RAM installed using a 64-bit version of Windows 10. We randomly generated the values of five inputs and three outputs using a uniform distribution with the minimum and maximum values from Table 5 for the total numbers of DMUs, {15, 50, 100, 150, 200, 300, 400, 500, 800, 1,000}. For comparison, *DEA Frontier*, a commercial DEA software package using an Excel spreadsheet as the primary file, is run for the generated DMUs on the same computer. The results are reported in Table 11. For S-DEA, we only measure the running time for generating the levels since finding ASAP requires another operation after finishing the level generation. Thus, the actual running time to find ASAP for S-DEA would be much greater than those times reported in Table 11.

As expected, the running times for MXN-EOE are almost negligible compared to those of DEA-based models. For 100 DMUs, it takes less than one second for MXN-EOE, while CE-DEA takes 88.2 seconds. Table 11 shows that the running times of DEA-based methods sharply increase when the number of DMUs increases. In fact, it takes more than one hour to get the results generated by CE-DEA for 1,000 DMUs, while it takes less than 7 seconds for the MXN-EOE. Thus, the proposed method is more suitable for evaluating and ranking a large set of DMUs in the context of DEA and big data.

Table 11: Comparison of running time

(Time unit: second)

n	15	50	100	150	200	300	400	500	800	1,000
MXN-EOE	0.2	0.4	0.6	1.0	1.1	1.5	1.9	2.3	4.2	6.5
SE-DEA	2.4	8.2	24.2	48.3	120.8	220.5	409.6	612.3	1356.4	2,347.2
CE-DEA	9.5	33.2	88.2	162.6	258.6	446.2	752.4	1,107.2	2416.4	4,135.2
S-DEA	12.5	39.0	90.2	204.4	395.3	801.4	1,360.2	1,856.4	3,472.2	6,053.7

n : Number of DMUs

SUMMARY AND CONCLUSIONS

As Big Data research becomes an essential area of operations analytics, DEA is evolving into data-enabled analytics and a data-oriented data science tool for various operational analytics. In fact, DEA-based methods enable decision-makers to measure the efficiency of DMUs and rank them based on efficiency scores (Ess). Many researchers show that Big Data can be applied to improve company productivity or efficiency and will be essential for enterprises to grow and achieve a competitive advantage. However, the DEA-based methods in the Big Data context require excessive running time to evaluate a large set of DMUs since these methods require optimization software to apply any DEA methods.

Ranking DMUs based on the Ess generated by the C-DEA method shows a significant drawback due to the self-evaluation principles. Several ranking methods based on the C-DEA have been

proposed, but no ranking method has been found to be either a universal or superior method for ranking DMUs. The authors, who developed various DEA-based ranking methods, have asserted that the absence of global assessment criteria makes evaluating all the presented methods reviewed by their papers impossible. They conclude that each method could be better than others according to the decision maker's preferences and evaluation objectives, depending on the evaluation's nature (see Aldamak and Zolfaghari, 2017). The severe weakness of the DEA-based approaches is their biased preference for specific inputs/outputs. All DMUs under evaluation can only use favorable inputs/outputs to boost their own efficiency scores, dropping the unfavorable inputs/outputs in the DEA context. In addition, if new DMUs are introduced to the evaluation set, DEA-based models must evaluate entire DMUs, including the newly introduced DMUs. Thus, DEA-based models are neither efficient in terms of computational time nor effective in terms of evaluating and ranking DMUs.

Machine learning is an approach to learning complex patterns from existing data and using these patterns to make predictions on unseen data. This paper proposes a machine learning algorithm (MLA) based on an efficiency-driven evaluation method with a max-normalization approach (MXN-EOE) to overcome the severe weaknesses of the DEA-based models, including the excessive running times required by the DEA models. In MXN, each input/output is divided by the maximum value for each column to remove the imbalance among the data. We apply the proposed method to the three well-known numerical examples, which several authors have considered to illustrate their own methods based on DEA models.

These numerical examples show that the rankings generated by the DEA-based methods show such a significant weakness, especially for the top-notch DMUs. The rankings generated by DEA-based models turn out to be neither robust nor consistent. Especially when the decision-makers are interested in the top efficient DMUs, applying DEA-based models would mislead them. In contrast to DEA-based methods, the MXN-EOE generates consistent and robust rankings and yields results very quickly. Table 18 shows that the proposed MXN-EOE takes less than 7 seconds to get the results with one thousand (1,000) DMUs, whereas the DEA software takes around 40, 68, and 100 minutes for SE-, CE-, and S-DEA. In addition, the proposed procedure does not need any optimization software. The results and observations through the numerical examples reveal that the MXN-EOE performs well in terms of running times and the robustness of generated rankings. Thus, we propose that the MXN_EOE method be considered an appropriate tool for evaluating a large set of DMUs in the Big Data context. Future research would be interesting and necessary to apply the ML-EOE method extensively in real-world DEA applications and the Big Data context.

REFERENCES

References are available upon request from Hong.

Socioeconomic Determinants of Banking Term Deposit Subscriptions: A Multivariate Analysis Approach

Prakarsha Mitra

Zarb School of Business, Hofstra University

Hempstead, New York, 11549

pmitra1@pride.hofstra.edu

Abstract

Consumer demand for products and services, including financial instruments like banking term deposits, is driven by the need to fulfill specific preferences and goals. This demand is influenced by various socioeconomic factors, which play a pivotal role in shaping consumer behavior. The **Theory of Reasoned Action** offers a valuable framework for understanding these influences, positing that consumer decisions are guided by rational tendencies associated with socioeconomic determinants such as income, education, employment status, gender, and social affiliations. This paper explores the relationship between these factors and consumer behavior toward subscribing to banking term deposits. Employing data from a Portuguese banking institution's telemarketing campaigns, the study integrates principal component analysis (PCA) and logistic regression to identify patterns and success predictors. The findings reveal that factors such as age, education, and history of previous contact positively influence subscription likelihood, while financial constraints, such as housing or personal loans, and excessive contact attempts decrease it. By examining these relationships, the study provides a robust framework for optimizing banking marketing strategies, showcasing how data-driven techniques can enhance predictive accuracy and campaign effectiveness. These insights contribute significantly to the understanding of consumer behavior in the financial sector, bridging theoretical foundations and practical applications in targeted marketing.

Introduction

Historically, a bank would need to undergo a merger or acquisition to achieve significant household growth. With targeted and data-driven marketing, that is no longer the case. The marketing function has evolved in many banks from primarily a communications function to a partner with lines of business to grow revenue. This evolution has been evident for years, but gained momentum during the pandemic and has become more essential as interest rates have squeezed industry profits (Gibson, 2024). Banks are always looking for ways to improve their

marketing efforts, and one important goal is to figure out who's likely to subscribe to term deposits. Term deposits are a common way for people to save and invest, but not everyone signs up for them. Term deposit investments usually carry short-term maturities ranging from one month to a few years and will have varying levels of required minimum deposits. The interest earned on a term deposit account is slightly higher than that paid on standard savings or interest-bearing checking accounts. The increased rate is because access to the money is limited for the timeframe of the term deposit (Vincent, 2023).

Advertisements attempt to portray a message that provokes certain attitudes and behaviors. Research has identified how advertisers can reach a specific audience. An example of consumer buying behavior is the psychoanalytical model. It considers how behaviors are affected by both the conscious and subconscious mind (Jisana, 2014). Utilizing the power of data-driven insights is not only beneficial but also essential. Financial institutions are under more pressure than ever to use data to inform their choices as they manage economic and market volatility. In the modern digital age, defined by a wealth of data and ongoing technological advancements, financial institutions must recognize the criticality of data-driven decision-making to their sustainable resilience and competition (Inamdar, 2024). Data and analytics can help you better understand your target audience by providing insights into their demographics, interests and behaviors. This information can be used to create more relevant and personalized marketing messages that are more likely to resonate with your audience (Akins, 2024). This paper explores the socioeconomic determinants influencing term deposit subscriptions, utilizing multivariate tools such as principal component analysis (PCA) and logistic regression to identify patterns and predictors of consumer behavior. By examining factors such as age, education, financial stability, and prior interactions, this study aims to provide actionable insights for improving targeted marketing efforts in the banking sector.

Literature Review

The literature on banking term deposits highlights the interplay between consumer behavior, socioeconomic determinants, and data-driven marketing strategies. Term deposits, characterized by their fixed maturity periods and competitive interest rates, are a vital financial product for both customers and banking institutions. Understanding the factors influencing their adoption is essential for optimizing marketing efforts and improving customer targeting. Term deposits are a low-risk investment option offering slightly higher interest rates than standard savings accounts in exchange for limited liquidity during the deposit term. They play a critical role in the banking ecosystem, allowing institutions to stabilize their funding while offering customers a secure investment avenue (Vincent, 2023). However, despite their benefits,

not all customers subscribe to term deposits, necessitating deeper research into the factors influencing consumer choices (Chen, 2022). Socioeconomic characteristics such as age, education, employment status, and financial stability significantly shape consumer decisions regarding term deposit subscriptions. Research indicates that older, more educated individuals are more likely to adopt term deposits due to their greater financial awareness and stability (Yi, 2024). Conversely, financial constraints such as housing or personal loans tend to reduce the likelihood of subscription, as these factors increase consumers' risk aversion and limit their capacity to invest in long-term financial products (Moro et al., 2014).

Behavioral models, including the psychoanalytical framework, provide insights into how both conscious and subconscious factors influence consumer decisions. This model emphasizes the role of emotions, past experiences, and social contexts in shaping behavior, which can be particularly relevant for financial decisions like term deposit subscriptions (Jisana, 2014). Additionally, the Theory of Reasoned Action highlights the importance of rational tendencies driven by socioeconomic determinants, providing a structured approach to understanding consumer behavior in financial contexts (Akins, 2024). The integration of data analytics in banking has revolutionized how financial institutions target and engage potential customers. Predictive modeling techniques such as principal component analysis (PCA) and logistic regression enable banks to identify patterns and success predictors within large datasets. These tools have proven effective in segmenting customers based on demographic, behavioral, and financial characteristics, thereby improving the precision of marketing campaigns (Moro et al., 2014; Akins, 2024).

Several studies underscore the importance of targeted marketing in driving term deposit subscriptions. For instance, prior successful interactions and longer call durations have been found to positively influence customer responses, demonstrating the value of high-quality engagement (Buddiga, 2022). Similarly, excessive contact attempts or poorly timed communication can have a negative impact, highlighting the need for a balanced approach to customer outreach (Guha et al., 2024). These findings underscore the critical role of optimizing communication strategies to enhance campaign effectiveness. This study builds on existing research by combining behavioral theories with advanced multivariate techniques to analyze the socioeconomic determinants of term deposit subscriptions. By employing PCA to uncover hidden patterns and logistic regression to quantify the effects of predictors, this research contributes to the growing body of knowledge on data-driven decision-making in financial marketing. Furthermore, the study provides actionable insights for banks to refine their marketing strategies, ensuring that they align with consumer behavior and preferences.

Methodology

The data analyzed in this study comes from direct marketing campaigns from a Portuguese banking institution through phone calls. The dataset is cited as DOI: 10.24432/C5K306, and its citation is as follows: Moro, S., Rita, P., & Cortez, P. (2014). Bank Marketing [Dataset]. UCI Machine Learning Repository. Retrieved from <https://doi.org/10.24432/C5K306>.

While the data is from Portuguese, the research can be extended to other countries applicable

The various methods of analysis in the paper are manifold. PCA was done to reduce dimensionality, decreasing similar variables to significant components. The Logistic Regression aims to model the probability of success based on a pattern identified and their predictors. Perhaps more important, factor analysis was conducted because it exposes the relationships that are not immediately apparent by reducing highly correlated variables into their root factors; thus, spotting those patterns that weigh customers' decisions.

Results

The principal component analysis produced seven dimensions of explanation in the data. The first dimension, Interaction History, has brought out the fact that the history of previous contact and successful past campaigns increase subscription rates. The second dimension, Contact Timing, showed that the level of engagement is heavily influenced by the timing and methods of communication. The third dimension, Demographics, showed that factors such as age and marital status shape preferences and increase the likelihood of subscription.

The fourth dimension, Communication Quality, indicates that long and meaningful conversations improve customer responses and likelihood of subscription. The fifth dimension, Financial Stability, indicates that higher account balances increase the likelihood of subscription, while loans reduce this likelihood. The sixth dimension, Campaign Activity, suggests that too much contact may actually decrease effectiveness and risk disengaging potential customers. The seventh dimension is Socioeconomic Background; it highlights that education and job profiles are crucial in decision-making and also depicts the trends of client segmentation.

A logistic regression representation was made based on improvements that involved the following effects, representing features that are the most influential to change the probability of the client subscribing. Amongst the listed factors, Positive Effects included the following: Age—

each additional year of age can raise the probability of subscribing by 1%; Higher Education increases the probability of subscribing by 17%. Additionally, longer call durations significantly enhance the likelihood of success, emphasizing the value of higher-quality communication. Moreover, more contacts—especially recent or successful interactions—contribute to an increased subscription probability of up to 21%.

However, Negative Impacts were also noticed. For example, clients with loans, whether housing or personal, are approximately 47% and 42% less likely to subscribe, respectively. If different methods of contact were used, some proved 44% less efficient, indicating that not all communication channels are equally effective. Lastly, high frequency of contact reduced the success rate by 9%, implying that excessive outreach might lead to customer fatigue or disengagement.

The factor analysis disclosed the bottom-line dimensions that are at play when determining client subscription behavior. Campaign Success History turned out to be an important factor, showing that past interactions are strong predictors of customer responsiveness. Demographics also play a vital role, with characteristics such as age and marital status contributing to or determining subscription likelihood.

Another critical factor that helps delineate the reasons in the data is Communication Style, reflecting the timing and methods of contact. Lastly, Financial Stability considers customer balances and loans as indicators of risk profiles that influence subscription responses. Together, these factors provide a deeper understanding of behavioral patterns in customer decision-making.

Discussion

Findings indicate that personal traits, finances, and communication are the three significant factors influencing the chances of subscription to term deposits. Older, well-established, and educated clients are more responsive toward these campaigns. Calling or following up too often results in negative subscription rates. Financial burdens such as mortgages and personal loans also harm subscription chances.

Longer calls are better than shorter or frequent ones, proving that content quality is a significant success factor.

Conclusion

Socioeconomic factors come to the front as predictors of term deposit subscription. While positive influences are exerted by age, education, and previous contacts, financial constraints and too frequent contacts diminish the rate of success. These findings provide actionable strategies for improving campaign performance and customer targeting.

It is interesting that more complex machine learning algorithms will, in the future, explore its incorporation for even greater refinement and assess dynamic personalization strategies in live marketing campaigns.

References

- Akins, A. (2024, December 10). Using data and analytics to improve marketing effectiveness. *ABA Banking Journal*. <https://bankingjournal.aba.com/2024/03/using-data-and-analytics-to-improve-marketing-effectiveness/>
- Analysis of predicting the success of the banking telemarketing campaigns by using machine learning techniques – *NORMA@NCI Library*. (n.d.). <https://norma.ncirl.ie/id/eprint/6337>
- Burnett-Nichols, H. (n.d.). Term Deposits: What They are and How They Work – *NerdWallet*. *NerdWallet Canada*. <https://www.nerdwallet.com/ca/banking/term-deposits-how-they-work>
- Buddiga, S. K. P. (2022). Improving Targeted Client Acquisition: Predictive Analysis of Retail Bank Direct Marketing Campaigns. *Jtipublishing.com*. <https://doi.org/10.93153/fqwgnb40>
- Chen, J. (2022, March 20). Term Deposit Definition. *Investopedia*. <https://www.investopedia.com/terms/t/termdeposit.asp>
- Finding the Best Techniques for Predicting Term Deposit Subscriptions (Case Study UCI Machine Learning Dataset). (2022, December 26). *IEEE Conference Publication | IEEE Xplore*. https://ieeexplore.ieee.org/abstract/document/10331379/?casa_token=RkrvqiazLkEAAAAA:xi1vRJBjuFKzZvo8ISxnmR1LE5sEWmXxVJIDZuzdBkvQ28RDy-xg6J54o0mv1i5sjQ4wuRtoo5hzmw
- Fitrianto, A., Muhamad, W. Z. a. W., & Susetyo, B. (2022). Development of direct marketing strategy for banking industry: The use of a Chi-squared Automatic Interaction Detector (CHAID)

in deposit subscription classification. *Journal of Socioeconomics and Development*, 5(1), 64. <https://doi.org/10.31328/jsed.v5i1.3420>

Gibson, M. (2024, September). The critical role of bank marketing in driving revenue growth. *ABA Banking Journal*. <https://bankingjournal.aba.com/2024/09/the-critical-role-of-bank-marketing-in-driving-revenue-growth/>

Guha, S., Dhar, M., Saha, B., Saha, N., & Department of Computer Science and Engineering, Guru Nanak Institute of Technology. (2024). A Machine Learning Approach to Identify Customers for Subscribing Term Deposit [Journal-article]. *ALOCHANA JOURNAL, VOLUME 13*(5, 2024).

Inamdar, N. (2024, December 30). Data-Driven Decision Making in Banking: unlocking powerful insights. *Lucent Innovation*. <https://www.lucentinnovation.com/blogs/it-insights/data-driven-decision-making-in-banking>

Moro, S., Cortez, P., & Rita, P. (2014). Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns. *Neural Computing and Applications*, 26(1), 131–139. <https://doi.org/10.1007/s00521-014-1703-0>

Prediction of Potential Customers for Term Deposit: Analysis Using Bagging with Bank Marketing Data. (2024, February 24). *IEEE Conference Publication | IEEE Xplore*. https://ieeexplore.ieee.org/abstract/document/10481968/?casa_token=7tv62uVRIMcAAAAA:rsolmgx_F38FiWsmcp66gNACd0hcxcWGzHIPfEQ-8lGg4vwJXksF4eMGaldLMPVj1O4aSPMbJFyk6Q

Singh, M., Dhanda, N., Farooqui, U. K., Gupta, K. K., & Verma, R. (2023). Prediction of Client Term Deposit Subscription Using Machine Learning. In *Lecture notes in electrical engineering* (pp. 83–93). https://doi.org/10.1007/978-981-99-2710-4_8

Tanvir, M. F., Hossain, M. M., & Jishan, M. A. (2024, October 28). Bayesian Regression for Predicting Subscription to Bank Term Deposits in Direct Marketing Campaigns. *arXiv.org*. <https://arxiv.org/abs/2410.21539>

Vincent. (2023, September 26). Term deposits classifications for bank marketing campaign. *Medium*. <https://medium.com/@yohanesvincentpangestu/term-deposits-classifications-for-bank-marketing-campaign-974f1a087e3b>

Yi, H. (2024). Predicting customer subscriptions to fixed-term deposit products based on machine learning approach. *Applied and Computational Engineering*, 50(1), 61–68. <https://doi.org/10.54254/2755-2721/50/20241218>

Zaki, A. M., Khodadadi, N., Lim, W. H., & Towfek, S. K. (2023). Predictive analytics and machine learning in direct marketing for anticipating bank term deposit subscriptions. *American Journal of Business and Operations Research*, 11(1), 79–88. <https://doi.org/10.54216/ajbor.110110>

Leveraging IoT for Environmental Sustainability: Enhancing Practices at Mercedes-Benz

Dr. Jessica Schwartz

71rexel71.schwartz@ucumberlands.edu

University of the Cumberlands, Williamsburg, KY, USA

Baker College, Owosso, MI, USA

APUS, Washington, DC, USA

Nexford University, Washington, DC, USA

Western Governors University, Salt Lake City, UT, USA

Dr. Patricia Jones (Presenting)

patricia.b.jones@wilmu.edu

Wilmington University, Wilmington, DE, USA

Baker College, Owosso, MI, USA

Dr. Lori Farr

lori.farr@ucumberlands.edu

University of the Cumberlands, Williamsburg, KY, USA

Western Governors University, Salt Lake City, UT, USA

Abstract

The Internet of Things (IoT) enables real-time data collection and automation, driving innovation in environmental monitoring, energy management, and sustainable practices. This paper explores Mercedes-Benz's application of IoT in environmental sustainability, highlighting its achievements and identifying areas for improvement. As a leader in luxury automotive production, Mercedes-Benz integrates IoT to reduce carbon emissions, optimize resource use, and enhance green initiatives, such as green charging. However, challenges remain in fully leveraging IoT for environmental monitoring and achieving real-time actionable insights. By examining empirical research and industry practices, this study proposes strategies to overcome technological, structural, and resource barriers. These include the integration of advanced IoT frameworks, interdisciplinary collaborations, and stakeholder engagement to foster innovation and sustainability. The paper emphasizes the potential of IoT to transform environmental monitoring in the automotive sector, aligning with Mercedes-Benz's circular economy goals. By addressing these gaps, the study offers a roadmap for optimizing IoT deployment, ensuring efficient energy management, and advancing the company's sustainability objectives.

IoT for Environmental Monitoring

The Internet of Things (IoT) refers to the interconnected network of physical devices, sensors, and systems that communicate and share data over the internet. By enabling real-time data collection and automation, IoT plays a crucial role in enhancing operational efficiency across various industries, from healthcare to manufacturing. Its growing adoption drives innovations in areas like environmental monitoring, smart cities, and energy management, shaping a more connected and sustainable future. Mercedes-Benz, a renowned German automotive manufacturer, is a leader in luxury and commercial vehicle production, with a global reputation for innovation and quality. Headquartered in Stuttgart, Germany, the company has embraced a forward-thinking approach toward sustainability, with the ambitious goal of achieving 50% electric vehicle (EV) capacity by 2030 to have a net carbon-neutral new vehicle fleet across all stages of the value chain by 2039, and have no impact on NO₂ urban area pollution by 2025 (Mercedes-Benz Group AG, 2023). The company's dedication to an electric and digital future is evident through its comprehensive sustainability initiatives, which focus on promoting inclusion, equity, and diversity alongside environmental responsibility (Dong & Ullah, 2023).

Mercedes-Benz is also actively exploring and incorporating cutting-edge technologies, such as the Internet of Things, to bolster its environmental sustainability efforts. The company's initiatives, such as green charging, aim to reduce carbon emissions and optimize the sustainability of its operations. They are also implementing Industry 4.0 principles, which rely heavily on IoT, by integrating them into production workflows to help optimize these workflows and use IoT sensors to track energy consumption using real-time data to help predict equipment replacement or maintenance (Mercedes-Benz Group AG, 2023). There remains room for improvement, especially in the integration of IoT technologies to enhance real-time environmental monitoring, energy efficiency, and resource optimization (Ullo & Sinha, 2020).

Through these advancements, Mercedes-Benz is poised to set new benchmarks in sustainable automotive production, leveraging IoT not only to reduce environmental impact but also to promote a circular economy. This paper examines the current practices of Mercedes-Benz in sustainability, the role of IoT in enhancing environmental monitoring, and how the company can further optimize its sustainability strategies through technology integration.

Background

Mercedes-Benz entails a strong commitment to sustainability that aims to empower employees and people in the world toward an electric and digital future while effectively promoting inclusion, equity, and diversity (Dong & Ullah, 2023). There are various areas where the company could improve its associated sustainability practices. This is significantly essential in the domain of environmental aspects that are used in the process of data analytics and monitoring. A major strong point of the company is the ability to embrace the Internet of Things technologies for practices in environmental monitoring. The company significantly focuses on green charging, which further dictates its commitment to effectively reducing the desired carbon emissions and properly optimizing environmental sustainability (Ullo & Sinha, 2020). Mercedes-Benz could further enhance efforts to promote sustainability by integrating IoT technologies based on environmental monitoring practices (Anandraj et al., 2021). Apart from the fact that the company demonstrates a proper commitment to optimizing renewable energy, room for improvement is required to help leverage IoT substantially in real time to monitor resource usage and vital environmental parameters adequate for optimization (Dong & Ullah, 2023). Mercedes-Benz can further enhance the deployment of capabilities that ensure actionable insights (Anandraj et al., 2021). This effectively facilitates a continuous environment and innovation for proper sustainability practices.

Regulatory Framework and Compliance

The integration of IoT technologies at Mercedes-Benz not only drives innovation and sustainability but also requires strict adherence to an evolving landscape of environmental and data privacy regulations. Mercedes-Benz must navigate complex regulatory requirements at both regional and international levels. Global and regional environmental regulations emphasize the need to reduce emissions, improve energy efficiency, and promote sustainable operations (Yunxia & Yuqing, 2025). Regulatory bodies in Europe and North America have established ambitious targets for reducing greenhouse gas emissions. IoT-enabled environmental monitoring systems can help Mercedes-Benz track real-time data on energy consumption and emissions, ensuring compliance with standards such as the European Union's CO₂ reduction policies and local air quality regulations (Yunxia & Yuqing, 2025).

The extensive data collection capabilities of IoT devices bring data privacy and security to the forefront, necessitating compliance with stringent data protection laws. As Mercedes-Benz operates within the European Union, GDPR mandates robust data protection measures, including secure and transparent data collection, storage, and processing by IoT devices. Compliance with GDPR requires clear data governance policies, consent mechanisms, and regular audits to prevent data breaches (What is GDPR, n.d.). In addition to GDPR, various regions have enacted specific data privacy laws (such as the California Consumer Privacy Act in the United States) that affect IoT data management. These regulations highlight the importance of user consent, data minimization, and the right to be forgotten, compelling companies to adopt privacy-by-design principles in their IoT deployments. Besides privacy regulations, frameworks such as the NIST Cybersecurity Framework and ISO/IEC 27001 guide the secure management of IoT data. By integrating these standards into their operational protocols, companies can mitigate risks associated with unauthorized access and cyber threats.

Implementation Challenges and Best Practices

The integration of IoT technologies for environmental monitoring and sustainability initiatives presents both significant opportunities and challenges. For companies like Mercedes-Benz, achieving the dual goals of technological advancement and robust cybersecurity requires addressing various technical, organizational, and financial barriers, all while safeguarding sensitive data and maintaining system integrity.

IoT devices inherently expand the network perimeter, exposing systems to vulnerabilities such as unauthorized access, data breaches, and the exploitation of insecure communication protocols. Common risks include inadequate encryption, unpatched firmware, and insecure data transmission methods. These vulnerabilities can lead to unauthorized data access or manipulation of critical environmental monitoring systems, thereby impacting operational efficiency and the reliability of sustainability data. Ullo and Sinha (2020) emphasize that cybersecurity frameworks are critical for maintaining data integrity and operational continuity in IoT systems. These frameworks include the NIST Cybersecurity Framework, ISO 27001, and IEC 62443.

The initial investment for deploying IoT devices, establishing secure networks, and ensuring ongoing maintenance can be substantial. However, studies by Alloui & Mourdi (2023) and Peter et al. (2023) have shown that the long-term benefits, such as improved operational efficiency, reduced downtime, and enhanced environmental performance, often justify upfront costs. Mercedes-Benz can create a resilient IoT ecosystem that drives environmental monitoring and sustainability while protecting critical data assets from emerging cybersecurity threats.

IoT for Environmental Performance

Arsawan et al. (2024) conducted an empirical literature study about how the Internet of Things, environmental protection, and sustainability are linked. The study analyzed 1294 articles by effectively analyzing the WOS and SCOPUS databases. The study realized the presence of six essential clusters that are important in guiding future research. The visual trend provided a proper guide that helps ensure effective collaboration between practitioners and researchers in numerous academic fields (Arsawan et al., 2024). This leads to scenarios that boost the enhancement in quantity and quality. The research protocols deployed in the first study entailed identifying keywords, organizing data as a research information system (RIS), retrieving databases, and carrying out multiple selections (Arsawan et al., 2024). Arsawan et al. (2024) conclude that technology and IoT are vital to ensuring a compelling performance when dealing with numerous transactions.

The study indicates that IoT, energy consumption, sustainable performance, digital technology, and transition model are essential factors that lead to the realization of IoT to ensure the optimization of environmental monitoring performance (Arsawan et al., 2024). Mercedes-Benz can deploy these fundamental mechanisms and concepts in its technological transformation to electric vehicles (Evs) (Arsawan et al., 2024). The deployment can, therefore, be an essential framework in the process that involves implementing essential constructs as a pathway to realizing optimized environmental performance (Ullo & Sinha, 2020). Mitigating barriers is vital, especially concerning systems that are unsupported in resistance to technology, have limited resources, have high costs, and are experiencing technological transformation (Arsawan et al., 2024). Furthermore, Mercedes-Benz can use the same concept to influence its partners, especially since the companies constantly explore the internet and technology to realize environmental legitimacy and a long-term strategic approach to planning (Jesse, 2018).

AI Deployment for IoT-Based Air Pollution Monitoring

Asha et al. (2022) deployed an Environmental Toxicology-based Air Pollution Monitoring System based on Artificial Intelligence (ETAPM-AIT) model. The study proposed the model to foster optimization of sensor array that is IoT-based in assessing humidity, temperature, and some pollutants that play a crucial role. The model developed in the study aims to successfully report the air quality status using a cloud-based server in real-time while being able to send an alarm if there are hazardous pollutants in the air. Even though IoT-based sensor networks have been discussed, the analysis has not been comprehensive, especially concerning the integration approach required for renewable energy sources in the monitoring framework (Ullo & Sinha, 2020). The study should have explored the potential impact associated with the incorporation of sustainable practices to ensure suitability in a Mercedes-Benz environment. Mercedes appreciated the decarbonization process, although the potential impact of incorporating sustainable practices in IoT-based environmental monitoring techniques has widely remained underexplored (Jesse, 2018).

Essential factors such as adaptability and scalability have not been effectively explored in the study, as per the ETAPM-AIT model, when considering the diverse settings of the environmental perspectives. Furthermore, the practical suitability and reliability of such practices at Mercedes Benz are hindered by relying too much on simulation analysis (Lukin et al., 2022). To handle such gaps, any research carried out in the future needs to prioritize the collaboration approaches to foster proper optimization and incorporation of sustainability principles (Izonin et al., 2023). An emphasis on stakeholder management is an essential pathway that studies can explore as a step closer to the adaption and acceptance of an IoT-based environmental monitoring system (Lukin et al., 2022). Considering interdisciplinary collaborations is key and should enhance essential data security measures (Jesse, 2018).

Value Chain and Software Ecosystems in Green IoT

Jesse (2018) assesses the effect of IoT practices, especially when considering the transition toward autonomous vehicles. Essential technical and market trends are considered, primarily based on the fact that there is a focus on aspects entailing industrial mechanisms that are influential following the future of transportation (Jesse, 2021). Integrating IoT in the automobile sector plays a key role as it properly facilitates narrowing the gap between autonomous driving and proper connectivity (Anandraj et al., 2021). This plays a pivotal role in revolutionizing the ecosystem in the designated industry.

Considering the initiatives of green computing, the purposeful mechanisms by Mercedes-Benz play a crucial role in boosting how IoT can be aligned. Adherence to renewable energy and green charging play a significant role based on the dedication to promoting practical environmental sustainability and minimizing the experienced carbon emissions (Anandraj et al., 2021). The study has played an influential role in ensuring the valuable realization of insights into opportunities and challenges correlated to autonomous vehicles and the desired opportunities (Ullo & Sinha, 2020). The emphasis on how vital proper issues can be addressed plays a significant role in the optimization of progress and mechanisms of stability (Anandraj et al., 2021).

IoT-Based Sensors for Environmental Monitoring

Sunny et al. (2020) deployed fundamental measurement principles when considering implementation mechanisms for demonstrating essential techniques for monitoring various environmental parameters such as humidity, temperature, and pressure. The study provides an exciting perspective involving using wireless sensor networks in harsh environmental conditions, emphasizing the optimization of environmental monitoring (Ullo et al., 2020). Mercedes-Benz can consider this aspect and perspective, especially considering its sustainable people plan (Izonin et al., 2023). This plays a fundamental role in exemplifying how IoT initiatives should be

aligned to optimize corporate sustainability strategies (Ullo et al., 2020). The commitment to ensure the empowerment of employees is based on initiatives such as recharge, re-skill, and re-shape (Ullo & Sinha, 2020). This is a crucial indicator and influencer of the dedication to social responsibility while properly highlighting the desired role to facilitate sustainable development for the desired workforce (Izonin et al., 2023).

Sunny et al. have explored the Commercial Off-The-Shell (COTS) elements associated with the need to complement the monitoring systems related to sustainability efforts by promoting energy-efficient and cost-effective solutions (Izonin et al., 2023). Interestingly, the research has properly emphasized simplicity and modularity while considering the synergy needed for a conducive and sustainable work environment (Ullo et al., 2020). Leveraging such mechanisms is essential for enhancing sustainability and social responsibility mechanisms (Jesse, 2018).

Mercedes-Benz, renowned for its commitment to sustainability, aligns closely with the principles outlined in the research on transitioning towards a Circular Economy (CE) within the electric vehicle (EV) battery ecosystem. The company's sustainability efforts include product innovation, supply chain management, and corporate responsibility initiatives (Li, 2022). Integrating insights from the research, Mercedes-Benz optimizes its sustainability efforts by focusing on several key areas, particularly leveraging IoT-based environmental monitoring for effective management.

Mercedes-Benz engages with policymakers and industry stakeholders in the regulatory component to advocate for supportive policies conducive to CE adoption. By collaborating with regulatory bodies, the company ensures that its operations align with evolving environmental standards and regulations, facilitating a smoother transition towards sustainable practices in battery manufacturing and usage (Arawan et al., 2024). In addressing structural barriers, Mercedes-Benz fosters collaboration across its value chain, working closely with suppliers,

partners, and customers to promote CE initiatives (Li, 2022). Through transparent communication and cooperation, the company streamlines processes and enhances the efficiency of material flows, contributing to optimizing resource utilization and waste reduction (Arawan et al., 2024).

Mercedes-Benz prioritizes consumer awareness and interest in CE transition, emphasizing the environmental benefits of EVs and sustainable battery solutions. Educating consumers and offering attractive incentives stimulates demand for eco-friendly products, driving market transformation towards CE-driven business models (Li, 2022). In terms of IoT-based environmental monitoring, Mercedes-Benz invests in advanced technology to track and optimize resource usage throughout its operations. The company gains real-time insights into energy consumption, emissions, and waste generation by deploying IoT sensors and data analytics platforms, enabling proactive decision-making to minimize environmental impact (Li, 2022).

In product development, Mercedes-Benz implements circular design principles, ensuring EV batteries are designed for longevity, reuse, and recycling. Through innovative design strategies and material selection, the company maximizes the lifecycle of its products while minimizing resource depletion and waste generation (Izonin et al., 2023). Mercedes-Benz's sustainability efforts align closely with the research findings on transitioning towards a Circular Economy within the EV battery ecosystem. By addressing regulatory, structural, and consumer-related barriers while leveraging IoT-based environmental monitoring and circular design principles, the company optimizes its operations for long-term environmental sustainability and business success.

IoT, Electric Vehicles, and Autonomous Vehicles

The automobile industry is experiencing a significant transformation, moving from self-driving cars to fully autonomous cars, from gas and diesel-powered vehicles to fully electric

vehicles (Evs), and vehicles with limited connectivity to vehicles that are fully connected to the digital network. Today's automobiles have been connected through links to technology like smartphones, roadside assistance, and real-time traffic alerts for some time now (Krasniqi & Hajrizi, 2016). Industry 4.0 brings a pivotal stage in the fourth industrial revolution to usher in changes, and IoT is a critical factor in enabling Industry 4.0 because it connects the physical device to the digital network. The automotive sector stands to significantly benefit from IoT adoption, especially in the supply chain, production logistics, manufacturing logistics, and customer feedback (Krasniqi & Hajrizi, 2016; Silva et al., 2020).

Industry 4.0 and IoT for Sustainability

Industry 4.0 (I4.0), referred to as the Fourth Industrial Revolution, is where manufacturing and industry are being transformed through a higher level of digitization and the integration of advanced and emerging technologies. This higher level of digitization results in flexible, efficient, sustainable, high-quality, low-cost production. The concept of Industry 4.0 is integrating a cyber-physical system (CPS) into manufacturing operations that use IoT in their industrial processes. Advanced manufacturing technologies include industrial IoT, which helps run these new industrial systems (Machado, 2020).

In I4.0, data is captured from the manufacturing process using IoT for production and optimization of the value chain (Bag et al., 2021). There are many environmental benefits, such as reduced water, electricity, and gas consumption during production. In addition, an organization's profit planning is impacted by eco-design and use. Eco-friendly design and innovation can help organizations build legitimacy and trust, meet the expectations of the public, and build solid credentials (Yu et al., 2021). Eco-design and eco-innovation are important strategies for innovation in manufacturing.

Eco-design and eco-innovation have different focuses. Eco-design focuses on the development of new, environmentally friendly technologies, products, and services. Eco-

innovation transforms operational manufacturing processes to reduce resource consumption, thus strengthening the production process efficiency (Alkaraan et al., 2024). Sustainability practices can improve an organization's performance (Song et al., 2018); therefore, Mercedes-Benz needs to continue to increase its efforts and resources to develop green, eco-friendly products.

Future Trends and Emerging Technologies

There are many future trends and emerging technologies that are driving innovation and helping to shape the future of IoT sustainability and green technology. These innovations include driving innovation in industry, optimizing resource use, helping mitigate climate change, and promoting circular economies. Future technologies include AI-powered IoT, 5G and low-power wide-area networks (LPWAN), sustainable smart cities, IOT-enhanced circular economy solutions, biodegradable and sustainable IoT devices, and green mobility and sustainable transportation.

AI-powered IoT is being used for sustainability in areas such as autonomous decision-making. IoT devices will make autonomous adjustments for carbon reduction, energy savings, and waste minimization by analyzing large amounts of data in real-time (Sami et al., 2024). AI will help businesses and governments be proactive in implementing sustainability measures by predicting environmental trends and working to reduce or optimize carbon emissions.

5G and Low-Power Wide-Area Networks (LPWAN) can be used to enhance IoT connectivity with ultra-fast communications with low latency for smart grids. LPWAN can support low-power, short-range connections that are ideal for water management, tracking pollution, and smart agriculture. These new network protocols can reduce battery consumption in smart sensors by optimizing data transfer rates (Ogbodo et al., 2022).

According to Belaid et al. (2024), sustainable smart cities are beginning to implement AI-driven traffic management where emissions and congestion can be minimized. IoT-renewable

energy grids are being implemented to improve a city's efficiency using sustainable technologies such as solar, wind, and hydroelectric power. IoT-enhanced circular economies can use IoT to implement smart waste management, IoT-based product lifecycle monitoring, and blockchain for supply chain transparency.

Biodegradable and sustainable IoT devices will use eco-friendly sensors and chips with biodegradable, non-toxic materials. IoT devices can be self-powered using solar, kinetic, and thermoelectric energy-harvesting technologies. Bioplastics and sustainable rare-earth materials will be used for recyclable electronics (Zahedi et al., 2024). According to Rodríguez-Rad et al. (2023), green mobility and sustainable transportation are already being used through IoT-integrated EV charging networks, connected autonomous vehicles (CAVs), and IoT-enabled shared mobility. This is evident with Mercedes-Benz being at the forefront of integrating IoT and sustainability into its operations.

Conclusion

The study explored the complex interaction between IoT, environmental monitoring, and sustainability, providing valuable insights for future research and business applications. The assessment of various literature has been significant in the study and has provided essential practices and mechanisms that entail technological mechanisms to promote sustainability. The literature has further provided a broader perspective of the mechanisms that need to be considered while assessing and comparing them to the already established sustainability practices at Mercedes-Benz. The company's commitment to sustainability, as demonstrated by its Sustainable People Plan and activities in electric vehicles and renewable energy, is consistent, and the issues are adequately covered in the literature review. The company might, however, need to optimize its processes in various sectors, notably incorporating IoT technologies into environmental monitoring systems. The study is crucial in emphasizing the need to combine renewable energy sources and optimize resource utilization via IoT-based

monitoring, which could help Mercedes-Benz's sustainability initiatives. Furthermore, the study underlines the need for interdisciplinary collaborations and stakeholder participation to ensure the practical usability and acceptability of IoT-enabled environmental monitoring solutions.

References

- Alkaraan, F., Elmarzouky, M., Hussainey, K., Venkatesh, V. G., Shi, Y., & Gulko, N. (2024). Reinforcing green business strategies with Industry 4.0 and governance towards sustainability: Natural-resource-based view and dynamic capability. *Business Strategy and the Environment*, 33(4), 3588–3606. <https://doi.org/10.1002/bse.3665>
- Allioui, H., & Mourdi, Y. (2023, Sept 22). Exploring the full potentials of IoT for better financial growth and stability: A comprehensive survey. *Sensors*, 23(19). <https://doi.org/10.3390/s23198015>
- Anandraj, A. P. S., Nandhini, P., Punitha, A. A. A., & Kaviyaraj, R. (2021, July). A new vehicular emergency model based on IoT. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)* (pp. 643-648). IEEE.
- Arsawan, I. W. E., Hariyanti, N. K. D., Azizah, A., Suryantini, N. P. S., & Darmayanti, N. P. A. (2024). Internet of Things towards environmental performance: A scientometrics and future research avenues. In *E3S Web of Conferences* (Vol. 501, p. 01011). EDP Sciences.
- Asha, P., Natrayan, L. B. T. J. R. R. G. S., Geetha, B. T., Beulah, J. R., Sumathy, R., Varalakshmi, G., & Neelakandan, S. (2022). IoT-enabled environmental toxicology for air pollution monitoring using AI techniques. *Environmental Research*, 205, 112574.
- Bag, S., Wood, L. C., Telukdarie, A., & Venkatesh, V. G. (2021). Application of Industry 4.0 tools to empower circular economy and achieving sustainability in supply chain operations. *Production Planning & Control*, 34, 918–940. <https://doi.org/10.1080/09537287.2021.1980902>
- Belaid, F., Ramdani, B., & Abdellatif, M. (2024). Can smart cities be sustainable? An emerging field of research. *Journal of Urban Technology*, 1–22. <https://doi.org/10.1080/10630732.2024.2352992>

- Dong, Z., & Ullah, S. (2023). Towards a green economy in China? Examining the impact of the Internet of Things and environmental regulation on green growth. *Sustainability*, 15(16), 12528.
- Izonin, I., Tkachenko, R., Krak, I., Berezsky, O., Shevchuk, I., & Shandilya, S. K. (2023). A cascade ensemble-learning model for the deployment at the edge: Case on missing IoT data recovery in environmental monitoring systems. *Frontiers in Environmental Science*.
- Jesse, N. (2018). Internet of Things and Big Data: The disruption of the value chain and the rise of new software ecosystems. *Ai & Society*, 33(2), 229-239.
- Khan, M. N., Tanvirahmedshuvo, Ontor, M. R. H., Khan N., & Rahman, A. (2024). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *International Journal for Multidisciplinary Research*, 6(1). 10.36948/ijfmr.2024.v06i01.22699
- Khan, Y., Su'ud, M. B. M., Alam, M. M., Ahmad, S. F., Ahmad, A. Y. A. B., & Khan, N. (2022, Dec 30). Application of Internet of Things (IoT) in sustainable supply chain management. *Sustainability* 2023, 15(1). <https://doi.org/10.3390/su15010694>
- Krasniqi, X., & Hajrizi, E. (2016). Use of IoT technology to drive the automotive industry from connected to full autonomous vehicles. *IFAC-PapersOnLine*, 49(29), 269-274.
- Li, X. (2022). An analysis of Mercedes-Benz marketing strategy. *Frontiers in Business, Economics and Management*, 6(1), 87-89.
- Lukin, E., Krajnović, A., & Bosna, J. (2022). Sustainability strategies and achieving SDGs: A comparative analysis of leading companies in the automotive industry. *Sustainability*, 14(7), 4000.
- Machado, C. G., Winroth, M. P., & Ribeiro da Silva, E. H. D. (2020). Sustainable manufacturing in Industry 4.0: An emerging research agenda. *International Journal of Production Research*, 58(5), 1462–1484. <https://doi.org/10.1080/00207543.2019.1652777>

Mercedes-Benz Group AG. (2023). *Sustainability report 2023*. Mercedes-Benz Group AG.

<https://group.mercedes-benz.com/sustainability/sustainability-report.html>

McKinsey Global Institute. (2015). *The Internet of Things: Mapping the value beyond the hype*.

McKinsey & Company.

Parviziomran, E., & Elliot, V. (2024). Barriers to circular economy: Insights from a small electric vehicle battery manufacturer. *Journal of Purchasing and Supply Management*, 100905.

Ogbodo, E. U., Abu-Mahfouz, A. M., & Kurien, A. M. (2022). A survey on 5G and LPWAN-IoT for improved smart cities and remote area applications: From the aspect of architecture and security. *Sensors (Basel, Switzerland)*, 22(16), 6313. <https://doi.org/10.3390/s22166313>

Peter, O., Pradhan, A., & Mbohwa, C. (2023). Industrial Internet of Things (IIoT): Opportunities, challenges and requirements in manufacturing business in emerging economies.

Procedia Computer Science, 217: 856-865. <https://doi.org/10.1016/j.procs.2022.12.282>

Rodríguez-Rad, C. J., Revilla-Camacho, M. Á., & Sánchez-Del-Río-Vázquez, M. E. (2023).

Exploring the intention to adopt sustainable mobility modes of transport among young university students. *International journal of environmental research and public health*, 20(4), 3196. <https://doi.org/10.3390/ijerph20043196>

Sami, J., Abdulrahman, L., Abdullah, R., Sami, T., Salim, B., & Abdulkareem, N. (2024). AI-powered sustainability management in enterprise systems based on cloud and web technology integrating IoT data for environmental impact reduction, 03, 156-176.

Silva, M., Signoretti, G., Silva, I., & Ferrari, P. (2020). Performance evaluation of a vehicular edge device for customer feedback in Industry 4.0. *Acta IMEKO*, 9(4), 88-,

https://doi.org/10.21014/acta_imeko.v9i4.719

Song, W., Ren, S., & Yu, J. (2018). Bridging the gap between corporate social responsibility and new green product success: The role of green organizational identity. *Business Strategy and the Environment*, 28(1), 88–97. <https://doi.org/10.1002/bse.2205>

- Sunny, A. I., Zhao, A., Li, L., & Sakiliba, S. K. (2020). Low-cost IoT-based sensor system: A case study on harsh environmental monitoring. *Sensors*, 21(1), 214.
- The ROI of IoT: The 7 benefits it can bring to your business. (n.d.) *World Sensing*.
- Ullo, S. L., & Sinha, G. R. (2020). Advances in smart environment monitoring systems using IoT and sensors. *Sensors*, 20(11). <https://doi.org/10.3390/s20113113>
- What is GDPR compliance? (n.d.). *Paloalto Networks*.
<https://www.paloaltonetworks.com/cyberpedia/gdpr-compliance>
- Yoon, D. (2024, May 2). Digital transformation outcomes: Why do companies need it to improve ROI. *KaiNexus*.
- Yu, Y., Zhang, J. Z., Cao, Y., & Kazancoglu, Y. (2021). Intelligent transformation of the manufacturing industry for Industry 4.0: Seizing financial benefits from supply chain relationship capital through enterprise green management. *Technological Forecasting and Social Change*, 172, 120999. <https://doi.org/10.1016/j.techfore.2021.120999>
- Yunxia, X., & Yuqing, N. (2025, Jan 22). The technological advancement of environmental regulations and their impact on energy efficiency and CO2 emissions. *Scientific Reports*, 15. <https://doi.org/10.1038/s41598-025-86581-z>
- Zahedi, A., Liyanapathirana, R., & Thiyagarajan, K. (2024). Biodegradable and renewable antennas for green IoT sensors: A review. *IEEE Access*, 1-1. 10.1109/ACCESS.2024.3515274.

Prompt Engineer - The New Programmer

Bay Arinze, PhD

Samir Shah, PhD

Department of Decision Sciences and MIS

LeBow College, Drexel University

January 15, 2025

Abstract

The advent of large language models (LLMs) has introduced a paradigm shift in software development, elevating the role of prompt engineering over traditional programming. This paper explores the evolving landscape where prompt engineering is poised to become a critical development skill, now complementing, but potentially supplanting conventional coding practices over time. While this transition is not yet fully realized, the argument is made that it is plausible in the medium to long term. The paper discusses the underlying reasons, mechanisms, and implications of this shift for organizational productivity, digital transformation, and system development methodologies. It also examines the role of modern-day Robotic Process Automation (RPA) in facilitating this transition.

Introduction

The rapid development of artificial intelligence (AI) and machine learning (ML) technologies, particularly large language models (LLMs) like OpenAI's GPT-3 and GPT-4, has begun to reshape various aspects of technology and industry. Among the most profound changes is the emergence of "prompt engineering" as a crucial skill set. Unlike traditional programming, which involves writing explicit code to define behavior, prompt engineering involves crafting queries or prompts to extract desired outputs and specifically, program code from AI models. This paper investigates whether prompt engineering could eventually replace traditional programming as the dominant development skill and examines the broader implications for the tech industry. Furthermore, the role of cloud computing and modern-day Robotic Process Automation (RPA) tools in enhancing productivity and accelerating digital transformation will be explored.

Evolution of the Programmer/Developer Role

The role of the programmer has evolved significantly since the inception of computing. In the early days, programming was a highly specialized skill. Early programmers worked directly with machine code and assembly language, which required an in-depth understanding of computer architecture and binary logic. The introduction of higher-level programming languages like FORTRAN, COBOL, and later C, allowed programmers to write more abstract code, making programming more accessible and less error prone. These early innovations laid the groundwork for the evolution of programming from a niche discipline into a cornerstone of modern technology (Ceruzzi, 2003; ACM, 2009).

By the 1970s and 1980s, efforts had been made to formalize the field of software engineering. Software development methodologies such as Waterfall and, later, Agile were developed to manage the growing complexity of software projects. During this period, the role of the programmer expanded to include software design, testing, and maintenance. The introduction of integrated development environments (IDEs) and version control systems further professionalized the field, enabling collaboration and more efficient workflows. This era also saw the emergence of software engineering as a recognized profession, with a greater emphasis on systematic approaches to design and development (Brooks, 1987; Pressman, 2014).

The explosion of the internet in the 1990s and 2000s brought about another significant shift. Programmers needed to learn new languages and frameworks to develop web applications, leading to the rise of full-stack development. The role of the developer became more multifaceted, requiring knowledge of both front-end and back-end technologies. Open-source software and collaborative platforms like GitHub transformed how software was developed and shared. This period also marked the beginning of the digital economy, with software development becoming central to business strategy and innovation (O'Reilly, 2005; Raymond, 2001).

The Age of AI, ML, and LLMs

In the past decade, AI and ML have begun to transform software development. Developers now need to understand data science, machine learning algorithms, and big data technologies. The role of the programmer has expanded to include training and deploying machine learning models, often requiring knowledge of specialized frameworks like TensorFlow and PyTorch. This transformation has led to the emergence of new interdisciplinary fields, blending traditional software engineering with data science and artificial intelligence (Jordan & Mitchell, 2015; LeCun, Bengio, & Hinton, 2015).

Large language models (LLMs) represent a significant leap in AI capabilities. These models, such as OpenAI's GPT-3 and GPT-4, are based on transformer architecture, which enables them to understand and generate human-like text. The development of LLMs began with simpler models like Google's BERT and progressed to more complex models capable of handling vast amounts of data and generating coherent and contextually relevant text.

The development of transformer models began with the introduction of the Transformer architecture by Vaswani et al. in 2017. This architecture revolutionized natural language processing (NLP) by enabling models to handle long-range dependencies and contextual information more effectively. The introduction of BERT (Bidirectional Encoder Representations

from Transformers) by Google in 2018 further demonstrated the power of transformers, setting the stage for subsequent advancements (Devlin et al., 2018).

OpenAI's Generative Pre-trained Transformer (GPT) models have been at the forefront of LLM development. GPT-2, released in 2019, demonstrated the ability to generate coherent text based on prompts. GPT-3, with 175 billion parameters, further pushed the boundaries, showing proficiency in tasks ranging from code generation to creative writing. The release of GPT-4 has continued this trend, with even more advanced capabilities and applications (Brown et al., 2020; Radford et al., 2019).

LLMs have demonstrated remarkable capabilities in various domains:

Natural Language Understanding: LLMs can understand and generate human-like text, making them useful for tasks such as language translation, summarization, and sentiment analysis. These models can process complex queries and provide detailed, contextually accurate responses.

Contextual Reasoning: These models can perform complex reasoning tasks, answering questions based on context and generating detailed explanations. This capability allows LLMs to engage in sophisticated dialogues and provide insights across a wide range of topics.

Code Generation: LLMs can assist in writing and debugging code, as seen with tools like GitHub Copilot, which uses GPT-3 to generate code snippets based on natural language prompts. This capability significantly reduces the time and effort required for coding tasks, enhancing developer productivity (Chen et al., 2021).

Relevant Use Cases

LLMs are now being applied in a wide range of industries and applications, such as:

Customer Support: Automated chatbots powered by LLMs provide customer support, handling inquiries and resolving issues without human intervention (Huang et al., 2022). These systems can understand and respond to customer queries in natural language, improving the efficiency and effectiveness of customer service operations.

Content Creation: LLMs are used to generate articles, marketing copy, and even creative writing, enhancing productivity in content-driven industries (Brown et al., 2020). By automating content generation, businesses can produce high-quality materials at scale, reducing costs and increasing output.

Healthcare: In the healthcare sector, LLMs assist in medical documentation, summarizing patient records, and even providing diagnostic suggestions (Lee et al., 2023). These applications improve the accuracy and efficiency of healthcare delivery, supporting clinicians in their decision-making processes.

Can Prompt Engineering Replace Traditional Programming?

Prompt engineering streamlines development processes by reducing the need for detailed coding knowledge. Developers now routinely leverage LLMs to generate and refine code, allowing them to focus more on high-level design and problem-solving. Studies also indicate that using LLMs can cut development time by up to 40% in certain scenarios (Lee et al., 2023). This reduction is achieved through the automation of routine coding tasks and the generation of high-quality code snippets.

With prompt engineering, individuals without extensive programming backgrounds can contribute to software development, democratizing access to technology (Williams et al., 2022). This accessibility allows more people to participate in the development process, fostering innovation and diversity in the tech industry.

Next, LLMs are highly adaptable, capable of performing diverse tasks across various domains. Prompt engineers can quickly switch contexts and apply LLMs to new problems without needing to learn new programming languages or frameworks. This flexibility allows organizations to respond more rapidly to changing business needs and technological advancements.

The Transition to Prompt Engineering

The transition from traditional programming to prompt engineering is driven by several factors:

Tooling and Infrastructure: The development of user-friendly interfaces and tools for interacting with LLMs is critical. Tools like OpenAI's playground and API, GitHub Copilot, and various IDE plugins enable prompt engineers to integrate LLMs seamlessly into their workflows. These tools provide the necessary infrastructure to leverage LLMs effectively, simplifying the prompt engineering process (Zhang et al., 2021).

Training and Education: As educational institutions and online platforms incorporate AI and ML into their curricula, a new generation of developers is emerging with skills in prompt engineering. Courses on prompt engineering, offered by platforms like Coursera and edX, are

gaining popularity. These educational initiatives are essential for equipping developers with the knowledge and skills needed to harness the full potential of LLMs (Ng, 2020).

Industry Adoption

Developers in all industries are already adopting prompt engineering practices. Those at the forefront include tech companies and startups, although adoption can now be described as pervasive.

Tech Companies: Leading tech companies like Microsoft, Google, and OpenAI are at the forefront of adopting LLMs, integrating them into their products and services to enhance functionality and user experience. These companies are setting benchmarks for the industry, demonstrating the value and potential of prompt engineering (Smith et al., 2021).

Startups: Many startups are leveraging LLMs to build innovative solutions in fields such as content creation, customer support, and personalized marketing. The flexibility of prompt engineering allows these companies to pivot quickly and adapt to market needs. This agility is particularly beneficial for startups, and particularly beneficial for startups operating in fast-paced, competitive environments (Williams et al., 2022; Zhang et al., 2021).

Implications for Productivity and Digital Transformation

The integration of LLMs and prompt engineering is already leading to significant productivity gains. LLMs automate routine coding tasks, allowing developers to focus on more complex and creative aspects of software development (Sadowski et al., 2023). This automation reduces the cognitive load on developers, enabling them to work more efficiently and effectively. Also, improved code quality results from generating code based on best practices and extensive training data, LLMs help to reduce bugs and improve code quality, thus enhancing overall project outcomes (Smith et al., 2023).

Role of Modern-Day RPA

Robotic Process Automation (RPA) tools like Make.com play a complementary, but pivotal role in enhancing productivity by automating repetitive tasks and processes. These tools integrate with LLMs to further streamline workflows and reduce manual intervention. RPA tools can automate tasks such as data entry, processing transactions, and generating reports, which frees up time for developers to focus on more strategic work (Xu et al., 2022). By handling mundane and

repetitive tasks, RPA tools increase overall efficiency and allow developers to concentrate on higher-value activities. By combining RPA with LLMs, organizations can create sophisticated automation solutions that handle complex tasks, such as natural language processing and decision-making, thereby significantly boosting productivity (Chen et al., 2023).

Accelerating Digital Transformation

The combination of LLMs and easy-to-use RPA platforms like Make.com have far reaching effects on digital transformation and agile methodologies. Prompt engineering is accelerating digital transformation in many organizations through rapid prototyping and scalability. The ability to quickly generate and test new ideas via rapid LLM-driven development allows companies to innovate faster and respond to changing market demands (Martin et al., 2021). This rapid prototyping capability is crucial for staying competitive in dynamic markets, enabling companies to experiment and iterate on new concepts with minimal delay.

Agile methodologies emphasize iterative development, collaboration, and flexibility. The integration of LLMs and prompt engineering can enhance these principles by enabling rapid prototyping and feedback loops. With prompt engineering, teams can quickly generate and test new ideas, accelerating the iterative process that is central to Agile practices. This capability allows for more frequent releases and faster response to user feedback, improving the overall agility of the development process. Furthermore, the ease of using LLMs means that non-developers can more easily participate in the development process. This inclusivity fosters greater collaboration within teams, bringing diverse perspectives and expertise into the project. As a result, the quality and innovation of the software produced are enhanced.

Moreover, LLMs can help streamline various Agile practices. For instance, in sprint planning, LLMs can assist in generating user stories, estimating task durations, and identifying potential risks. During daily stand-ups, they can provide summaries of progress, highlight blockers, and suggest solutions based on historical data. In retrospective meetings, LLMs can analyze team performance and suggest actionable improvements. These enhancements make Agile processes more efficient and effective, allowing teams to focus on high-priority tasks and strategic decisions (Sutherland & Schwaber, 2014).

Discussion and Conclusion

The transition from traditional programming to prompt engineering represents a momentous shift in software development. While it is not yet fully realized, the potential benefits in terms of efficiency, productivity, and accessibility suggest that prompt engineering could become the primary development skill in the medium to long term. The integration of LLMs with agile methodologies, cloud computing, and modern-day RPA tools will further facilitate this transition, driving digital transformation and innovation. By embracing these new technologies and methodologies, organizations can position themselves for success in an increasingly AI-driven world.

While the potential for prompt engineering to replace traditional programming is significant, some limitations and concerns must be addressed:

Model Limitations: Current LLMs are not infallible and can produce incorrect or biased outputs. Prompt engineers must be vigilant in verifying and refining responses (Zhang et al., 2023). Ensuring the accuracy and fairness of AI-generated content is critical for maintaining trust and reliability in AI systems.

Ethical Concerns: The use of LLMs raises ethical issues related to data privacy, bias, and accountability (Anderson et al., 2022). Addressing these concerns requires robust governance frameworks and ethical guidelines to guide the development and deployment of AI technologies.

The future of software development is likely to see a continued shift towards AI-assisted development. As LLMs become more powerful and accessible, their adoption across industries will likely increase, further driving the demand for prompt engineering skills. This trend will be supported by ongoing advancements in AI technology and the growing recognition of its potential to transform various sectors (Bishop, 2020).

The roles of software developers and prompt engineers will continue to evolve, with traditional programming skills becoming increasingly complemented by prompt engineering expertise. This evolution will require continuous learning and adaptation, as developers integrate AI tools into their workflows and expand their skill sets (Jordan & Mitchell, 2015).

References

1. Anderson, M., et al. (2022). Ethical Implications of AI and ML in Software Development. *Journal of Ethical AI Research*.
2. Bishop, C. M. (2020). *Pattern Recognition and Machine Learning*. Springer.
3. Brooks, F. P. (1987). *No Silver Bullet: Essence and Accidents of Software Engineering*. IEEE Computer Society.
4. Brown, T., et al. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*.
5. Ceruzzi, P. E. (2003). *A History of Modern Computing*. MIT Press.
6. Chen, M., et al. (2021). Evaluating Large Language Models Trained on Code. *arXiv preprint arXiv:2107.03374*.
7. Chen, X., et al. (2023). Integrating RPA with AI: A New Frontier in Automation. *Automation Journal*.
8. Davies, S., et al. (2023). Cloud Computing and AI: A Symbiotic Relationship. *International Journal of Cloud Computing*.
9. Devlin, J., et al. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.
10. Huang, J., et al. (2022). The Impact of AI on Customer Support: A Case Study. *Journal of Business and Technology*.
11. Johnson, L., et al. (2022). Scaling AI with Cloud Technologies. *Journal of Cloud Engineering*.
12. Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
13. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
14. Lee, K., et al. (2023). Accelerating Software Development with AI. *IEEE Transactions on Software Engineering*.
15. Martin, R., et al. (2021). Agile Development and AI Integration. *Journal of Agile Development*.
16. Ng, A. (2020). *Machine Learning Yearning*. Self-published.
17. O'Reilly, T. (2005). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Communications of the ACM*, 49(4), 13-14.
18. Pressman, R. S. (2014). *Software Engineering: A Practitioner's Approach*. McGraw-Hill Education.
19. Radford, A., et al. (2019). Language Models are Unsupervised Multitask Learners. *OpenAI Blog*.
20. Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'Reilly Media.

21. Sadowski, C., et al. (2023). Code Generation with AI: A Productivity Study. *Software Engineering Journal*.
22. Smith, J., et al. (2021). Enhancing Collaboration in Software Teams with AI. *Journal of Collaborative Computing*.
23. Sutherland, J., & Schwaber, K. (2014). *The Scrum Guide*. Scrum.org.
24. Williams, P., et al. (2022). Democratizing Software Development through AI. *Journal of Technological Innovation*.
25. Xu, Y., et al. (2022). Enhancing Productivity with RPA: A Case Study. *Journal of Automation and Robotics*.
26. Zhang, X., et al. (2023). Challenges in AI-Based Software Development. *Journal of Computer Science*.

Quantum Machine Learning Approaches in Malware Analysis:

An Extended Review

Philip Shen

School of Business and Technology
Marymount University
Arlington, VA, USA

<https://orcid.org/0000-0002-6308-2248>

Jennifer Ferreras-Perez

School of Business and Technology
Marymount University
Arlington, VA, USA

<https://orcid.org/0009-0003-0972-1824>

Abstract

Modern malware's escalating sophistication and evasiveness require advanced analytical solutions beyond traditional methods. While Artificial Intelligence (AI) provides some relief, its high processing power demands often result in bottlenecks, particularly when handling large datasets. This paper investigates the potential of quantum machine learning to revolutionize malware analysis. Quantum algorithms have shown the capability to outperform classical counterparts significantly, with advancements in quantum support vector machines, quantum convolutional neural networks, and n-gram analysis showing promise in malware recognition. The paper also addresses the challenges inherent in the early stages of quantum computing, such as the unique data formatting requirements, the problem of decoherence and the susceptibility of quantum machines to environmental disturbances, which demand sophisticated error correction mechanisms. Despite these challenges, quantum machine learning is poised to transform cybersecurity. By significantly enhancing the speed and efficiency of malware analysis, it offers security operations centers a crucial edge, potentially shifting the balance in favor of defense. This study highlights the critical need for ongoing research and collaboration in quantum machine learning to strengthen our cybersecurity infrastructure.

Keywords: Artificial intelligence (AI), quantum computing, malware analysis, n-gram

Introduction

The landscape of cybersecurity is in a constant state of flux, with the development of detection evading malware representing a significant challenge. Malware authors continuously refine their techniques to bypass traditional detection mechanisms (Akhtar & Feng, 2022; Or-Meir et al., 2019). Initial malware detection relied on a signature-based approach. This method was effective against known threats but fell short when confronted with new, unknown malware variants, leading to the evolution of detection strategies (Or-Meir et al., 2019; Ucci et al., 2019). Signature-based methods gave way to heuristic and behavioral detection techniques (Akhtar & Feng, 2022; Kara, 2019). Malware authors quickly adapted, developing techniques to evade even these more dynamic approaches.

In this arms race, artificial intelligence (AI) has emerged as a powerful ally for cybersecurity experts. AI can analyze vast datasets and identify complex patterns indicative of malware (Harini et al., 2022). With the increase in the volume and complexity of malware, classical computing infrastructures are reaching their limits. Kalinin and Krundyshev estimated that computational bottlenecks become a significant challenge when handling datasets exceeding 10^6 samples (Kalinin & Krundyshev, 2023).

The ability to analyze and characterize malware in real time could provide cybersecurity teams with an optimized response, minimizing damage and curtailing the spread of the infection (Akhtar & Feng, 2022; Dutta et al., 2022). The most promising solution to the problem of modern malware analysis is to harness the power of quantum computing.

Methodology

The methodology for this focused literature review is designed to answer the pivotal question: “How can quantum machine learning enhance the efficiency and accuracy of malware analysis in comparison to traditional methods?” To address this, a comprehensive search strategy targeting peer-reviewed articles, conference proceedings, and reputable industry reports published from 2015 to 2023. No relevant articles were found before 2019. Databases/search engines such as IEEE Xplore, ACM Digital Library, and Google Scholar were used. The snowball method was also used. Keywords included “quantum machine learning”, “quantum malware detection”, “quantum malware analysis”, “quantum algorithms” and “quantum computing”. Papers which met the search criteria were retrieved for further review. Non-English language papers were excluded from further analysis. The full text was read if the publication met the initial screen. All of the relevant articles were utilized in this review.

Malware Analysis

Malware analysis has evolved into a critical component of cybersecurity, tasked with dissecting malicious software to understand its origin, functionality, and potential impact on affected systems (Dutta et al., 2022). The traditional cornerstone of malware analysis was the signature-based approach, where security systems scanned for known patterns within code to identify threats. This method, while straightforward, is only effective against previously encountered and cataloged malware. Its inherent limitation is the inability to detect novel or sophisticated threats that do not match any existing signatures (Dutta et al., 2022; Kara, 2019). Consequently, the cybersecurity community has shifted towards heuristic and behavioral analysis, which examines the actions of a program to infer malicious intent (Akhtar & Feng, 2022; Kara, 2019). These methods allow for the identification of malware based on how it

behaves rather than how it looks, enabling the detection of polymorphic or previously unknown malware strains.

Despite the advancements in heuristic techniques, the surging volume and complexity of malware present a significant challenge, often overwhelming traditional analytical methods. In response, Artificial Intelligence (AI) has been increasingly incorporated into malware analysis. AI, particularly machine learning algorithms, can learn from and adapt to new malware features, identifying threats without relying on static databases of known malware signatures (Harini et al., 2022). However, the effectiveness of AI is heavily reliant on the availability of vast amounts of data and substantial computational resources to process and analyze this data efficiently (Akhtar & Feng, 2022; Dutta et al., 2022). As malware becomes more adept at evading detection through obfuscation and other evasion techniques, the demand for more powerful and intelligent analysis tools grows. The requirement for extensive processing capabilities and the ability to learn from evolving data patterns underscores the need for innovative approaches in malware analysis that can keep pace with the ingenuity of cyber threats (Harini et al., 2022).

Neural Networks

Neural networks have been one of the more successful machine learning approaches which have been described in the literature. Machine learning algorithms have been used for malware analysis, but authors often present the algorithm that they created and comparative studies are lacking. The current state of development of neural networks is illustrated by the work of Parihar et al. (Parihar et al., 2022). These authors enhanced the neural network approach by combining three neural networks to classify malware. The S-DCNN model is a sophisticated ensemble tailored to classify malware binaries into their respective families with impressive accuracy. By converting binary files into image-based formats, the model leverages the image recognition prowess of neural networks, circumventing the need for traditional and labor-intensive methods like feature engineering or reverse engineering. It integrates the power of three preeminent convoluted neural network architectures—ResNet50, Xception, and EfficientNet-B4—each contributing its unique strengths to the ensemble. These are then further refined through fine-tuning on malware-specific datasets, enhancing their capability as feature extractors for malware images. The combinatorial approach by Parihar et al is an innovative use of neural networks. Unless a much better neural network algorithm is created it is somewhat difficult to see how much more progress can be achieved by using a neural network approach running on a classical computer.

Quantum Computing

To appreciate the potential revolution quantum computing brings, it is essential to grasp the fundamental distinctions between it and classical computing. Classical computers, the backbone of technological advancement for decades, operate on bits as the primary data unit. These bits are binary, meaning they exist in one of two possible states: a zero or a one (Outeiral et al., 2021). In contrast, quantum computing introduces the quantum bit, or qubit, which leverages the principles of quantum mechanics (Gill et al., 2022). Unlike classical bits, qubits possess a characteristic known as superposition, allowing them to exist in a state that represents both zero and one simultaneously. This characteristic significantly expands computational possibilities, enabling quantum computers to process complex calculations at speeds unattainable by their classical counterparts.

Moreover, when multiple qubits are entangled, another quintessential quantum

phenomenon, the state of one entangled qubit instantly influences the state of others, regardless of the physical distance between them (Gill et al., 2022). This entanglement enables quantum computers to perform intricate computations with exponential growth in processing capability, as a system of N qubits can theoretically handle information on the order of 2^N , compared to classical systems that manage N bits at a time. Additionally, while classical computing hinges on the binary operation of the *NOT* gate to flip bit states, quantum computing utilizes a more complex set of operations (Taha, 2023; Tehrani et al., 2023). Quantum gates, which manipulate qubits, are represented by matrices in linear algebra, where a gate affecting N qubits is described by a $2^N \times 2^N$ matrix. This matrix-based operation enables quantum computers to undertake many calculations simultaneously, underscoring their ability to process and analyze vast amounts of information much more efficiently than classical computers, (Akhtar & Feng, 2022; Mezquita et al., 2021).

By leveraging the unique aspects of quantum mechanics quantum computers have the potential to dissect and analyze malicious software with unprecedented speed and depth. This advancement promises a more detailed and accurate extraction of information from malware samples and the potential to significantly reduce the time and resources currently required for such analyses.

Quantum Simulators

In quantum malware analysis research, the scarcity of accessible quantum computers necessitates an alternative approach for experimentation and development. Quantum simulators have emerged as a crucial tool in this context, standing in for the far less available quantum hardware (Ciaramella et al., 2023; Taha, 2023). These simulators are sophisticated pieces of software that mimic the behavior of a quantum computer, allowing researchers to model and test quantum algorithms and their applications for malware detection and analysis. The capabilities and features of quantum simulators, however, are not uniform. They vary widely depending on the specific simulator in question (Tehrani et al., 2023). This diversity in functionality and fidelity can result in considerable discrepancies in the performance outcomes of quantum algorithms, as underlined by de Avila et al. in their comparative study of seven different quantum simulators (de Avila et al., 2020).

The variability among quantum simulators poses a significant challenge when benchmarking the speed and efficiency of various quantum malware analysis methodologies. Since each simulator may emulate quantum computation with different levels of precision and efficiency, the results obtained from one simulator might not directly translate to another, let alone to an actual quantum computer (Barrue & Quertier, 2023; Taha, 2023). This makes it difficult to establish standardized comparisons. Consequently, while quantum simulators are invaluable for advancing research without widespread quantum computing resources, they also introduce complexity to the evaluation of progress within the field. Researchers must navigate these disparities, acknowledging the limitations while optimizing available simulation tools to push the boundaries of quantum malware analysis.

Quantum Computing for Malware Analysis

The unparalleled computational capabilities of quantum computers offer significant advancements over traditional computing methods. An overview of quantum algorithms by Outeiral et al. showed that they can outpace classical machine learning algorithms by a considerable margin, which is critical when analyzing large datasets that classical computers

struggle with due to limitations in processing power and memory (Outeiral et al., 2021). Moreover, combining machine learning with quantum computing has given rise to novel approaches for malware detection and classification. Techniques such as quantum support vector machines, quantum convolutional neural networks, and the use of n-grams are at the forefront of this research (Gill et al., 2022; Mezquita et al., 2021). These methods not only promise faster processing times but also more efficient utilization of computational resources.

Hybrid Quantum Neural Networks

The Hybrid Quantum Neural Network (HQNN) model proposed by Iqbal et al. was devised as a countermeasure to the escalating frequency of cyber-attacks (Iqbal et al., 2023). The HQNN is a blend of classical neural networks (CNN) and the capabilities of quantum computing. The architecture of the HQNN includes a classical encoder that receives input data and an encoding layer that translates the data into a quantum-readable format. This is followed by a quantum decoder that processes information through multiple quantum layers (Iqbal et al., 2023). The authors used a quantum simulator to test their model. By incorporating quantum principles, the HQNN model exhibits a remarkable increase in speed during both the training and testing phases when benchmarked against traditional CNNs. The result is a network that not only utilizes CPU resources more efficiently, achieving near-total utilization compared to the suboptimal performance of classical models, but also promises a significant reduction in response time to cyber threats. The downside of the authors' model is that it was not as accurate as the CNN model.

Suryotrisongko and Musashi created a hybrid quantum machine learning (QML) model for malware analysis, and they used a botnet dataset to test their model (Suryotrisongko & Musashi, 2022). Their testing was done on a quantum simulator. The quantum layer was sandwiched between classical layers in their model. They created six different QMLs. The accuracy of their QMLs was on average not as good as compared to their classical machine learning model. The highest accuracy was around 95%. The currently available quantum computers have been described as “Noisy Intermediate Scale Quantum” (NISQ) technology due to sensitivity to environmental factors. The most interesting aspect of this study is that they introduced additional noise into their QMLs. The authors found that the accuracy was dependent on which QML and noise combination was being tested. Again, the overall performance was not as good as their classical machine learning model but comparable to their models without the addition of the noise component.

Quantum Analysis on Quantum Computers

Tehrani, et al. looked at the performance of five classifier algorithms – VQC, PegasosQSVC, QSVC, SamplerQNN and Estimator QNN (Tehrani et al., 2023). These programs were run on quantum computers created by three different companies – IonQ, Rigetti and Quantinuum. The authors used a botnet dataset for their study. These algorithms were only about 40 to 60% accurate. The number of qubits available to the researchers was not stated. The accuracy was also limited because the capacity of the quantum computers was such that only 100 data points could be processed. The authors did note that the current quantum hardware suffers from errors and the machines available to them are NISQ machines. The accuracy of the five algorithms was improved (45-90%) when run on quantum simulators. This is because quantum simulators are not prone to the environmental sources of error that afflict quantum computers. The authors also developed a quantum algorithm which they were able to show performed well

in a quantum simulator. The algorithm the authors developed was a variant of the quantum Hoeffding tree classifier algorithm (QHTC). They altered the QHTC to run in batch mode instead of looping. Their version of QHTC was 100% accurate after three runs on the Rigetti simulator. While these results are promising and better than the results reported in many other papers, it remains to be seen how feasible quantum computing for malware analysis is when quantum computers with greater capacity and error correction become more available.

N-Grams and Quantum Analysis

Integrating N-grams and quantum computing redefines the approach toward malware analysis. N-grams, essentially sequences of 'n' items extracted from a text or speech sample, have long been a staple in natural language processing and statistical models. Their primary function is to predict and analyze data patterns, a capability crucial in identifying and categorizing characteristics of malicious software. As such, N-grams are valuable tools for detecting and classifying various malware families (Allgood & Nicholas, 2022; Taha, 2023). When combined with the advanced capabilities of quantum computing, this traditional approach opens new avenues for enhancing malware identification and analysis.

In a novel approach, Allgood and Nicholas proposed the utilization of N-grams in tandem with quantum computing to boost the efficiency of malware identification (Allgood & Nicholas, 2022). They leveraged Grover's Algorithm, a quantum algorithm known for rapidly searching unsorted databases, to expedite the association of N-grams to their corresponding hashes. This quantum method markedly streamlines the detection process, presenting a more efficient alternative to traditional techniques. The model was tested on a quantum simulator, demonstrating the feasibility and effectiveness of this innovative approach in the context of cybersecurity.

Challenges of Quantum Malware Analysis

Despite the potential benefits, applying quantum computing in malware analysis also has significant challenges. Current research, predominantly reliant on quantum computer simulators, has not yet been widely validated on physical quantum machines, which are constrained by the number of available qubits. There are, however, promising developments, such as the deployment in 2023 of a 127-qubit quantum computer at a university campus (Ciaramella et al., 2023; Malatino, 2023). The gap between lab-based research and practical commercial applications remains significant, and substantial advancements in quantum hardware are required before quantum computing can be routinely applied to malware analysis.

Data formatting and conversion from classical to quantum-compatible formats pose another hurdle. Quantum neural networks require data in a quantum state, which necessitates conversion algorithms that can be complex and time-consuming to implement (Outeiral et al., 2021). Additionally, interpreting the outputs from quantum computations introduces another layer of complexity, potentially hindering the seamless integration of quantum computing into existing malware analysis workflows.

Another significant drawback is the physical limitations and error susceptibility of quantum computers. The currently available quantum computers are NISQ machines and some of the issues were noted by Tehrani et al. These machines are susceptible to environmental factors such as temperature fluctuations and electromagnetic interference, which can introduce computational errors (Mezquita et al., 2021). This phenomenon is called decoherence. Similar issues affect quantum gates. Consequently, quantum computations are prone to high error rates

which can compromise the reliability of malware analysis outcomes. As quantum systems scale, the challenge of error correction becomes more daunting. The overhead for error-correcting algorithms and the need to dedicate more precious qubits to error correction could negate the computational speed advantages that quantum machines promise (Outeiral et al., 2021).

These challenges underscore the need for ongoing research and development in quantum computing. While the theoretical advantages of malware analysis are clear, the practical implementation is still hindered by several technological and logistical obstacles. Addressing these drawbacks is crucial for realizing the full potential of quantum computing in cybersecurity and malware analysis.

Discussion

As classical computers increasingly grapple with the complexity and stealth of modern malware, quantum computing emerges as a transformative solution, offering a significant leap in processing power and efficiency (Ciaramella et al., 2023; Gill et al., 2022; Taha, 2023). This capability is especially crucial as the field contends with the limitations of traditional artificial intelligence in detecting sophisticated malware. With its proficiency in managing complex operations and vast datasets, quantum computing is poised to redefine malware classification. This same power, however, could be leveraged by malicious actors. The biggest concern has been the potential for quantum computing to make existing encryption standards obsolete. Quantum malware does not yet exist. Some initial thoughts about this future possibility were noted by Dr. Stolfo of Columbia University (Stolfo, 2023). The current state of quantum computing has been called the NISQ era, and this topic was comprehensively reviewed by Lau et al. (Lau et al., 2022). These authors stated that the capacity and accuracy of current quantum computers is limited by the noise from environmental factors causing a quantum system to lose its quantum properties i.e., decoherence. There are three leading technical approaches for creating a quantum computer – superconducting qubits, trapped ions and photonics. According to the authors these platforms are all NISQ devices. The physics and engineering of these three approaches are very different and limit, for different reasons, the ability to increase the number of qubits and to decrease the error rate including decoherence problems. There is a great deal of research activity addressing these problems in order to create a fault tolerant quantum computer. Lau et al. estimate that it may take a decade or more to create such a computer. Charles Babbage created his analytical engine in 1837. The first electronic classical computer was ENIAC which was built during World War II. It took decades of development before classical computers reached their current state of utility. It was not until 1998 that the first (two qubit) quantum computer was built (Holton, 2023). Quantum computers are still in the early stages of development and like other new technologies the full potential of quantum computing remains to be realized.

Conclusion

This paper has explored the emerging role of quantum machine learning in malware analysis, a field increasingly challenged by the sophistication and evasiveness of modern cyber threats. Quantum machine learning presents a promising solution. Quantum algorithms, capable of outpacing classical counterparts by substantial margins, leverage techniques such as quantum support vector machines, quantum convolutional neural networks, and n-gram analysis for malware recognition. Quantum machine learning in cybersecurity has the potential to accelerate malware analysis. This offers security

operations centers a critical advantage in responding to malware attacks.

Realizing this potential requires overcoming significant challenges. These include the limited availability of quantum hardware, the complexity of data conversion for quantum analysis, and the high error rates in current quantum computations. As quantum technology advances, addressing these challenges is critical for integrating quantum computing into practical malware analysis.

In conclusion, the promise of quantum computing in the field of malware analysis offers a speed and efficiency improvement unattainable by classical computing. Continued investment in this field could profoundly impact our digital security landscape.

References

- Akhtar, M. S., & Feng, T. (2022). Malware Analysis and Detection Using Machine Learning Algorithms. *Symmetry*, 14(11). <https://doi.org/10.3390/sym14112304>
- Allgood, N., & Nicholas, C. (2022). A Quantum Algorithm to Locate Unknown Hashgrams. *Proceedings of the Future Technologies Conference*, 273–285. <https://arxiv.org/pdf/2005.02911.pdf>
- Barrue, G., & Quertier, T. (2023). *Quantum Machine Learning for Malware Classification*. <https://arxiv.org/pdf/2305.09674.pdf>
- Ciaramella, G., Martinelli, F., Mercaldo, F., & Santone, A. (2023). Exploring Quantum Machine Learning for Explainable Malware Detection. *2023 International Joint Conference on Neural Networks (IJCNN)*, 1–6. <https://doi.org/10.1109/IJCNN54540.2023.10191964>
- de Avila, A. B., Reiser, R. H. S., Pilla, M. L., & Yamin, A. C. (2020). State-of-the-art quantum computing simulators: Features, optimizations, and improvements for D-GM. *Neurocomputing*, 393, 223–233. <https://doi.org/https://doi.org/10.1016/j.neucom.2019.01.118>
- Dutta, N., Jadav, N., Tanwar, S., Deva Sarma, H., & Pricop, E. (2022). *Introduction to Malware Analysis* (pp. 129–141). https://doi.org/10.1007/978-981-16-6597-4_7
- Gill, S. S., Kumar, A., Singh, H., Singh, M., Kaur, K., Usman, M., & Buyya, R. (2022). Quantum computing: A taxonomy, systematic review and future directions. *Software: Practice and Experience*, 52(1), 66–114. <https://doi.org/https://doi.org/10.1002/spe.3039>
- Harini, S., Ravikumar, A., & Keshwani, N. (2022). Malware Prediction Analysis Using AI Techniques with the Effective Preprocessing and Dimensionality Reduction. In *Innovative Data Communication Technologies and Application Proceedings of ICIDCA 2021* (pp. 153–169). https://doi.org/10.1007/978-981-16-7167-8_12
- Holton, W. C. (2023, October 29). *Quantum computer*. Encyclopedia Britannica. www.britannica.com/technology/quantum-computer
- Iqbal, A., Alam, M., Javaid, N., Kazmi, N., Ahmad, F., & Urooj, A. (2023). Hybrid Quantum Neural Network Approach for Rapid Response to Cyber Attacks. *Journal of Computing & Biomedical Informatics*, 4(2). <https://www.jcbi.org/index.php/Main/article/view/138/87>
- Kalinin, M., & Krundyshev, V. (2023). Security intrusion detection using quantum machine learning techniques. *Journal of Computer Virology and Hacking Techniques*, 19(1), 125–136. <https://doi.org/10.1007/s11416-022-00435-0>
- Kara, I. (2019). A basic malware analysis method. *Computer Fraud & Security*, 2019(6), 11–19. [https://doi.org/https://doi.org/10.1016/S1361-3723\(19\)30064-8](https://doi.org/https://doi.org/10.1016/S1361-3723(19)30064-8)
- Lau, J. W. Z., Lim, K. H., Shrotriya, H., & Kwek, L. C. (2022). NISQ computing: where are we and where do we go? *AAPPS Bulletin*, 32(1), 27. <https://doi.org/10.1007/s43673-022-00058-z>
- Malatino, K. (2023, June 28). *Rensselaer Polytechnic Institute Plans to Deploy First IBM Quantum System One on a University Campus*. <https://news.rpi.edu/content/2023/06/28/rensselaer-polytechnic-institute-plans-deploy-first-ibm-quantum-system-one>
- Mezquita, Y., Alonso, R. S., Casado-Vara, R., Prieto, J., & Corchado, J. M. (2021). A Review of k-NN Algorithm Based on Classical and Quantum Machine Learning. In S. Rodríguez González & et al. (Eds.), *Distributed Computing and Artificial Intelligence, Special Sessions, 17th International Conference* (pp. 189–198). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-030-53829-3_20
- Or-Meir, O., Nissim, N., Elovici, Y., & Rokach, L. (2019). Dynamic Malware Analysis in the Modern Era—A State of the Art Survey. *ACM Comput. Surv.*, 52(5). <https://doi.org/10.1145/3329786>
- Outeiral, C., Strahm, M., Shi, J., Morris, G. M., Benjamin, S. C., & Deane, C. M. (2021). The prospects of quantum computing in computational molecular biology. *WIREs Computational*

- Molecular Science*, 11(1). <https://doi.org/10.1002/wcms.1481>
- Parihar, A. S., Kumar, S., & Khosla, S. (2022). S-DCNN: stacked deep convolutional neural networks for malware classification. *Multimedia Tools and Applications*, 81(21), 30997–31015. <https://doi.org/10.1007/s11042-022-12615-7>
- Stolfo, S. (2023, February 10). *Some Initial Thoughts about Quantum Malware*. <https://doi.org/https://doi.org/10.7916/ddhw-tp67>
- Suryotrisongko, H., & Musashi, Y. (2022). Evaluating hybrid quantum-classical deep learning for cybersecurity botnet DGA detection. *Procedia Computer Science*, 197, 223–229. <https://doi.org/https://doi.org/10.1016/j.procs.2021.12.135>
- Taha, B. N. (2023). An Investigation of Quantum and Parallel Computing Effects on Malware Families Classification. *Journal of Applied Science and Technology Trends*, 4(2), 72–83. <https://doi.org/10.38094/jastt42153>
- Tehrani, M., Sultanow, E., Buchanan, W. J., Amir, M., Jeschke, A., Chow, R., & Lemoudden, M. (2023). Enabling Quantum Cybersecurity Analytics in Botnet Detection: Stable Architecture and Speed-up through Tree Algorithms. *ArXiv E-Prints*, arXiv:2306.13727. <https://doi.org/10.48550/arXiv.2306.13727>
- Ucci, D., Aniello, L., & Baldoni, R. (2019). Survey of machine learning techniques for malware analysis. *Computers & Security*, 81, 123–147. <https://doi.org/https://doi.org/10.1016/j.cose.2018.11.001>

Navigating Uncertainty: Technology and Artificial Intelligence (AI) Sentiment in ESG Reports Pre- and Post-Pandemic^{*}

Jyothsna G. Harithsa[†]
SUNY Geneseo

Kofi Arhin[‡]
Lehigh University

February 2025

Abstract

Technology will help solve existing climate, societal, and governance-related problems in the next century, thus highlighting the crucial role of tech firms' leadership. This paper examines how leadership of tech firms listed in the S&P 500 communicate the use of advanced technology in their sustainability efforts, via environmental, social, and governance (ESG) reports. We use the impact of the COVID-19 pandemic, which saw an accelerated use of technology and also emphasized the need for more robust ESG efforts to solve global issues, as the setting for our study. Analyzing 571 ESG reports of 92 technology firms spanning the years 2008-2024, we find that firms were more strategic in their tone about technology post-covid compared to pre-covid years. Tech-firms performed better post-COVID, regardless of ESG reporting practices, with statistically significant improvements in profitability and firm value. However, surprisingly, firms with published ESG reports and with a communicated emphasis on technology did not see a significant increase in profitability. Additionally, positive communication about technology correlated with higher firm performance, but strategic vision and uncertain tones did not show significant or consistent results. Barring the significant impact of positive tones regarding technology in ESG reports on firm performance, the findings overall raise questions about the tangible efficacy of such disclosures.

Keywords: Social Responsibility, Corporate Governance, Natural Language Processing
JEL Classification: M14, M15, G30

^{*} All errors are our own. First Version: January 2025.

[†]Corresponding Author, State University of New York at Geneseo, Geneseo, NY, USA. Email: jharithsa@geneseo.edu

[‡]Lehigh University, Bethlehem, PA, USA. Email: kofi.arhin@lehigh.edu

1 Introduction

Communicating sustainability efforts and social responsibility practices is essential for organizations to thrive (Cohen et al., 2023; Lee et al., 2023). These efforts are primarily measured using environmental, social, and governance (ESG) metrics (Atkins et al., 2023). Meanwhile, existing studies have shown that artificial intelligence (AI) and other emerging technologies can help organizations achieve their ESG goals. For example, Zhang and Yang (2024) posit that AI adoption can significantly improve organizational ESG performance. Similarly, Burnaev et al. (2023) reveal that the adoption of AI can improve organizational ESG performance in the areas of clean water, sanitation, and forest preservation, among others. Transparent ESG communication via reports signals accountability and a genuine commitment to addressing global challenges, which can enhance an organization's reputation (Yu et al., 2018).¹

Moreover, the COVID-19 pandemic heightened the urgency for robust efforts to address ESG-related concerns,² as it exposed vulnerabilities in global supply chains, labor practices, and public health systems, prompting organizations to address these challenges more transparently and sustainably (Broadstock et al., 2021; Gao and Geng, 2024). Thus, organizations face mounting pressure from stakeholders to improve ESG performance (Candio, 2024). Understanding how sentiments around AI and technology are articulated in ESG reports can provide useful insights for theory and practice in this regard. Existing studies have shown that sentiments in organizational reports can provide valuable insights into how organizations perceive and prioritize resources (Berger et al., 2024; Chen et al., 2021). Yet, there remains limited research on the discourse surrounding AI and technology use for ESG goals.

In this study, we aim to uncover trends related to how organizations position technology and AI tools as strategic initiatives, especially in light of the transformational shifts brought

¹See the PwC Global Investor Survey: <https://www.pwc.com/gx/en/news-room/press-releases/2023/pwc-2023-global-investor-survey.html>

²Over 50% of investors surveyed by J.P. Morgan Research believed back in 2020 that the pandemic will act as a positive catalyst for ESG investing, see <https://www.jpmorgan.com/insights/global-research/esg/covid-19-esg-investing>

about by the COVID-19 pandemic. We seek to fill this gap by systematically examining how sentiments on technology and AI have evolved over time. Since strategy and uncertainty are drivers of organizational culture (Owusu and Gupta, 2024), we examine how COVID-19 impacted technology and AI sentiments related to strategy and uncertainty.

2 Data and Variable Definitions

To achieve the goal of the study, we analyze Chief Executive Officer (CEO) letters in organizational ESG reports (N=571) from 92 unique technology firms that have been components of the S&P 500 index between the years 2008-2024. Some firms drop off the list during the sample years and some join the index in the middle of the study period, hence we use firm CEO letters only for the years that each firm is a constituent of the S&P 500 index. We use the Way Back Machine, a website that archives past web pages, to track the index constituents over the years. In rare instances (like in the case of Meta Platforms, Inc.) the letter from management in the annual Corporate Responsibility Report is penned by a member of senior management who is not the CEO.

We manually collect the letters addressed to readers, written by the Chief Executive Officer in most cases (or in some rare cases, by a senior member of the C-Suite), in the annual Corporate Responsibility (or, ESG Reports) for each firm.³ We use CEO reports because communication from C-Suite to shareholders is representative of the firm's attitudes and cultural values (Amernic et al., 2010), which can have an impact on firm performance (Grennan, 2020). Such communicated corporate ethical values and CSR-based rhetoric can represent the overall ESG values of the firm.

We treat each letter as a firm-year data point and extract CEO sentiments from them. We identify the mentions of technology-related words in the letters, see Table A1 for the full list of tech-words. We then obtain the positive sentiment in relation to the use of technology, by

³The reports are either readily available on each firm's website, or archived on www.ResponsibilityReports.com

employing RoBERTa, a deep learning-based language processing model (Liu, 2019).⁴ Next, we measure uncertainty and strategy sentiments concerning technology in the CEO letters using a dictionary provided by Loughran and McDonald (2011).⁵ The variables extracted from the letters, namely Tech Word Use, Tech Word Sentiment, Certainty Score, Strategy Score, and World Events, are described in Table A2 in the Appendix.

We obtain control variables and firm performance variables from Compustat. We use the return on equity (ROE), return on assets (ROA), Tobin's Q, Net Profit Margin, and Operating Margin to proxy for firm value and firm performance, while we control for the size of the firm, the firm's age, liquidity, and leverage capabilities.

Table 1 exhibits the summary statistics of our dependent variables and independent variables for our overall sample, for the pre-covid years (2008-2019), and for the post-covid years (2020-2024). Panel A shows the CEO letter characteristics, which are derived using Natural Language Processing tools to extract meaning out of textual content based on tone. CEO letters became more wordy post-covid. The characteristics of the CEO letters are derived using Natural Language Processing tools to extract meaning out of textual content based on tone. CEOs exhibited a higher awareness of the role of technology in ESG, evidenced by the increase in the number of tech words per document as well as the percentage of tech words. While positivity pertaining to contextual technology words decreased, the contextual uncertainty and strategy and overall uncertainty and strategy increased post-covid compared to pre-covid years. Panel B represents our firm performance variables; all variables barring operating margin saw an increase post-covid. Finally, Panel C contains our control variables. Interestingly, firms lowered their current ratio and increased their leverage post-covid, indicating an increase in risky behaviors.

⁴See <https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest>

⁵ Uncertainty and strategy-related sentiments measured from written CEO discourse are representative drivers of organizational risk and corporate culture (Owusu and Gupta, 2024; Garcia et al., 2024).

3 Empirical Strategy

To test our hypotheses, we use an Ordinary Least Squares (OLS) regression. First, we determine the impact of the COVID-19 pandemic on the tone of tech-related words in CEO letters.

$$CEOTone_{it} = \beta_0 + \beta_1 postcovid_{it} + \beta_2 X_{it} + \delta + \epsilon_{it} \quad (1)$$

where $CEOTone_{it}$, the independent variable, is the positive score, the overall uncertainty and strategy of the CEO letters, or the contextual uncertainty or strategy score assigned to the two sentences preceding and following the mention of technology-related buzzwords in the CEO responsibility letters for firm i for year $= t$. $postcovid_{it}$ is a dummy variable equal to one for the years following the coronavirus pandemic (2020-2024), and zero for the years preceding. X_{it} represents the control variates used in our study, namely firm size, age, liquidity, and firm leverage. Standard errors are clustered at the firm-level, and our regressions include firm, year, and state fixed effects, which are represented in the equation by δ . Our fixed effects capture the impact of annual trends and time-invariant heterogeneities such as state regulations, and other macroeconomic variables.

Next, we study the impact of the COVID-19 pandemic on tech firm performance.

$$Performance_{it} = \beta_0 + \beta_1 postcovid_{it} + \beta_2 X_{it} + \delta + \epsilon_{it} \quad (2)$$

where $Performance_{it}$, our variable of interest, is a measure for firm value, proxied by the return on equity, return on assets, Tobin's Q, net profit margin, and net operating margin. $Postcovid_{it}$, X_{it} , and δ are as previously defined.

Finally, we study the impact of drastic changes to leadership tone post-COVID-19 pandemic on tech firm performance.

$$\text{Performance}_{it} = \beta_0 + \beta_1 \text{postcovid}_{it} * T \text{ oneChange}_{it} + \beta_2 T \text{ oneChange}_{it} + \beta_3 \text{postcovid}_{it} + \beta_4 X_{it} + \delta + \epsilon_{it} \quad (3)$$

where Performance_{it} , postcovid_{it} , X_{it} , and δ are the same as they have been defined in Equations 2-3 above and $T \text{ oneChange}_{it}$ is a dummy variable equal to 1 for firms in the higher median of largest change in tone relating to positivity, uncertainty, or strategy post-covid compared to pre-covid, and 0 for firms in the lower median representing tone change.

4 Results

We employ regression analysis to analyze the impact of the coronavirus pandemic on the textual content published by CEOs via responsibility reports. First, post-covid, we find that tech firms significantly increase their usage of tech-related buzzwords, exhibiting an increase of 2.73 words compared to the pre-pandemic era. They also mention tech words more scaled by total word count, as shown by the 0.4% increase in the percentage of tech words, see Table 2. Next, we show the results for the sentiments expressed by CEOs in their corporate letters in Table 3. For the entire period of study of 2008-2024 (see Panel A), except for contextual strategy (that is, strategy related to tech words), there was no significant change in firm-communicated sentiments. However, interestingly, for the period closest to the pandemic (2015-2024), there was a statistically significant (between 0.2%-4.7%) increase in overall uncertainty, and both overall and contextual strategy.

Technology firms, irrespective of whether they published a CSR report or not, performed significantly better post-covid compared to pre-covid, as evidenced by a higher ROA, Profit Margins, and Operating Margin (between +3.6% to 7.3% in profitability) and higher firm value (as indicated by the increase in Tobin's Q), see Panel A of Table 4. Surprisingly, the results were not significant for firms with published CSR or ESG reports or for firms

that mentioned futuristic technologies in their ESG reports, see Panels B and C of Table 3, barring a statistically significant increase in Tobin's Q for firms with published ESG reports post-covid, compared to the 2015-2019 era.

Firms that were more positive about technology in their communication exhibited higher firm performance (between a 3.6% - 7% increase in profitability) post-COVID compared to the pre-COVID era, as shown in Panel B of Table 5. Interestingly, firms using uncertain tones (overall and contextually) exhibited better firm performance overall (conflicting results for Tobin's Q), with a 5.8%-14.8% increase in performance. However, firms that expressed higher strategic vision exhibited significant decline in firm performance (between a 2.3% - 5.76% decrease in profitability). The results are similar when we focus our analysis on the years 2015-2024, see Table 6.

5 Conclusion

The study's findings highlight the need for organizations to prioritize technological innovation and adaptability as a key part of their business strategy. The superior post-COVID performance of technology firms, irrespective of their ESG reporting practices, underscores the resilience and strategic advantages of technology-driven organizations during crises. The lack of significant performance differences raises questions about the tangible impact of such disclosures. This suggests that while ESG communication is valuable for transparency, it may not directly translate into measurable improvements in financial performance. We encourage managers and policymakers to reevaluate the alignment between ESG efforts, technological investments, and financial outcomes, ensuring that initiatives are actionable and effective, rather than purely symbolic.

References

Amernic, J., Craig, R., and Tourish, D. (2010). Measuring and assessing tone at the top using annual report CEO letters. The Institute of Chartered Accountants of Scotland.

- Atkins, J., Doni, F., Gasperini, A., Artuso, S., La Torre, I., and Sorrentino, L. (2023). Exploring the effectiveness of sustainability measurement: which esg metrics will survive covid-19? *Journal of Business Ethics*, 185(3):629–646.
- Berger, A. N., Kim, H. H., and Ma, X. (2024). Bank sentiment and liquidity hoarding. *Contemporary Accounting Research*, 41(3):1513–1542.
- Broadstock, D. C., Chan, K., Cheng, L. T., and Wang, X. (2021). The role of esg performance during times of financial crisis: Evidence from covid-19 in china. *Finance research letters*, 38:101716.
- Burnaev, E., Mironov, E., Shpilman, A., Mironenko, M., and Katalevsky, D. (2023). Practical ai cases for solving esg challenges. *Sustainability*, 15(17):12731.
- Candio, P. (2024). Evolution of esg and corporate performance in health care. In *Sustainability and Corporate Performance in Health Care: ESG Implications for the European Industry*, pages 67–83. Springer.
- Chen, W., Wu, H., and Zhang, L. (2021). Terrorist attacks, managerial sentiment, and corporate disclosures. *The Accounting Review*, 96(3):165–190.
- Cohen, S., Kadach, I., Ormazabal, G., and Reichelstein, S. (2023). Executive compensation tied to esg performance: International evidence. *Journal of Accounting Research*, 61(3):805–853.
- Gao, M. and Geng, X. (2024). The role of esg performance during times of covid-19 pandemic. *Scientific reports*, 14(1):2553.
- García, R. E., Harithsa, J., and Owusu, A. (2024). Adding stress in banking: Stress tests and risk-taking sentiments. *Journal of Corporate Finance*, page 102596.
- Grennan, J. (2020). Communicating culture consistently: Evidence from banks. Available at SSRN 3350645.
- Lee, M. T., Raschke, R. L., and Krishen, A. S. (2023). Understanding esg scores and firm performance: Are high-performing firms e, s, and g-balanced? *Technological Forecasting and Social Change*, 195:122779.
- Liu, Y. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 364.
- Loughran, T. and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65.
- Owusu, A. and Gupta, A. (2024). Identifying the risk culture of banks using machine learning. *International Journal of Managerial Finance*, 20(2):377–405.
- Yu, E. P.Y., Guo, C. Q., and Luu, B. V. (2018). Environmental, social and governance transparency and firm value. *Business Strategy and the Environment*, 27(7):987–1004.

Zhang, C. and Yang, J. (2024). Artificial intelligence and corporate esg performance. *Inter-national Review of Economics & Finance*, 96:103713.

6 Tables

Table 1: Summary statistics.

The definitions for five of our variables, extracted using Natural Language Processing (NLP) tools, is shown in the table below. We extract positive sentiments using the Roberta text classification tool. For the uncertainty and strategy sentiments, we first measure these sentiments overall across the entire document, and second, only for two sentences following and preceding tech-related words, thus capturing overall uncertainty and strategy, as well as tech-related uncertainty and strategy.

	All Firms			Pre-Covid			Post-Covid		
	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
Panel A: Document Characteristics									
Words	226	707.2566	398.307	94	676.6277	463.382	132	729.0682	344.816
Tech Count	226	3.2655	3.713	94	3.1596	4.482	132	3.3409	3.066
Tech Percentage	226	0.0046	0.004	94	0.0043	0.003	132	0.0047	0.004
Positive Score: Tech Words	226	0.8007	0.160	94	0.8148	0.152	132	0.7907	0.165
Uncertainty: Overall	226	0.0059	0.004	94	0.0057	0.004	132	0.0061	0.005
Strategy: Overall	226	0.2134	0.044	94	0.2061	0.042	132	0.2186	0.045
Uncertainty: Tech Words	226	0.0053	0.007	94	0.0052	0.007	132	0.0055	0.006
Strategy: Tech Words	226	0.2077	0.067	94	0.1885	0.060	132	0.2213	0.069
Panel B: Firm Performance Variables									
ROE	226	0.0246	0.183	94	0.0166	0.256	132	0.0302	0.106
ROA	226	0.1026	0.075	94	0.1003	0.073	132	0.1043	0.076
Tobin's Q	226	3.9043	2.850	94	3.1601	2.099	132	4.4343	3.185
Net Profit Margin	226	0.1719	0.131	94	0.1647	0.122	132	0.1771	0.137
Operating Margin	226	0.2190	0.147	94	0.2197	0.147	132	0.2184	0.147
Panel C: Control Variables									
LogAssets	226	10.1556	1.178	94	10.4021	1.276	132	9.9800	1.074
age	226	45.6858	25.989	94	47.5000	26.635	132	44.3939	25.542
CurrentRatio	226	1.8839	1.260	94	2.0501	1.387	132	1.7656	1.152
leverage	226	0.7151	5.257	94	0.5340	1.772	132	0.8441	6.723

Table 2: The impact of COVID-19 on text content.

The table reports the impact of the COVID-19 pandemic on CEO letter characteristics, such as total word and character count, count of technology buzzwords, and percentage of tech words. Post covid is a dummy variable equal to 1 for firms that publish CSR reports the period after the coronavirus pandemic (starting from the year 2020) and zero for the years prior. Control variables include firm characteristics such as firm age, firm size, and firm liquidity. All regressions include firm, state, and year fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) Words	(2) Chars	(3) Tech	(4) Tech_count	(5) Tech_percentage
post_covid	197.050 (126.081)	1326.422 (811.825)	0.522*** (0.195)	2.725** (1.183)	0.004*** (0.001)
Observations	501	501	501	501	501
Adjusted R ²	0.322	0.317	0.298	0.284	0.370

Table 3: The impact of COVID-19 on text content.

The table reports the impact of the COVID-19 pandemic on CEO letter characteristics, such as total word and character count, count of technology buzzwords, and percentage of tech words. Post covid is a dummy variable equal to 1 for firms that publish CSR reports the period after the coronavirus pandemic (starting from the year 2020) and zero for the years prior. Control variables include firm characteristics such as firm age, firm size, and firm liquidity. All regressions include firm, state, and year fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) Contextual Positivity	(2) Overall Uncertainty	(3) Overall Strategy	(4) Contextual Uncertainty	(5) Contextual Strategy
Panel A: All years					
post-covid	0.033 (0.163)	-0.000 (0.002)	0.020 (0.017)	-0.003 (0.007)	0.120** (0.060)
Observations	226	501	501	238	238
Adjusted R ²	0.100	0.104	0.433	0.102	0.313
Panel A: Years 2015-2024					
post-covid	-0.003 (0.071)	0.002* (0.001)	0.023** (0.010)	0.002 (0.003)	0.047* (0.025)
Observations	211	431	431	223	223
Adjusted R ²	0.112	0.112	0.458	0.110	0.317

Table 4: The impact of COVID-19 on letter sentiments.

The table reports the impact of the COVID-19 pandemic on sentiments of CEOs in their corporate responsibility letters. Panel A shows the post-pandemic performance of all tech firms present in the S&P 500 index compared to pre-pandemic. Panel B reports the same but only for five years preceding and following the pandemic. Post covid is a dummy variable equal to 1 for firms that publish CSR reports the period after the coronavirus pandemic (starting from the year 2020) and zero for the years prior. The dependent variables are tech-related positivity, uncertainty, strategy, and overall uncertainty and strategy. Control variables include firm characteristics such as firm age, firm size, and firm liquidity. All regressions include firm, state, and year fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ROE	(2) ROA	(3) Tobin's Q	(4) Net Profit Margin	(5) Operating Margin
Panel A: All firms					
PostCovid	0.060 (0.052)	0.066*** (0.015)	2.130*** (0.304)	0.073*** (0.025)	0.036*** (0.014)
Observations	1146	1147	1146	1147	1147
Adjusted R ²	0.157	0.325	0.571	0.369	0.750
Panel A2: Post-2015					
PostCovid	-0.015 (0.059)	0.028** (0.012)	1.258*** (0.296)	0.017 (0.022)	0.017 (0.012)
Observations	696	697	696	697	697
Adjusted R ²	0.189	0.505	0.680	0.469	0.795
Panel B: Firms with Published ESG Reports					
post_covid	-0.030 (0.164)	0.003 (0.029)	0.965 (0.609)	-0.026 (0.054)	-0.003 (0.030)
Observations	500	501	500	501	501
Adjusted R ²	0.110	0.448	0.752	0.454	0.764
Panel B2: Post-2015					
post_covid	-0.026 (0.113)	0.024 (0.017)	1.873*** (0.399)	0.002 (0.035)	0.012 (0.019)
Observations	430	431	430	431	431
Adjusted R ²	0.127	0.514	0.767	0.451	0.763
Panel C: Firms with Tech Mentions in ESG Reports					
post_covid	-0.040 (0.170)	-0.015 (0.053)	0.582 (1.597)	-0.005 (0.107)	-0.006 (0.068)
Observations	238	238	238	238	238
Adjusted R ²	0.243	0.610	0.721	0.636	0.811
Panel C2: Post-2015					
post-covid	0.047 (0.073)	0.049** (0.022)	1.685** (0.680)	0.037 (0.046)	0.018 (0.029)
Observations	223	223	223	223	223
Adjusted R ²	0.243	0.623	0.723	0.639	0.813

Table 5: The impact of COVID-19 on firm performance.

The table reports the impact of the COVID-19 pandemic on firm value and firm performance. Panel A shows the post-pandemic performance of all tech firms present in the S&P 500 index compared to pre-pandemic. Panel B reports the same but only for those firms that published ESG reports, while Panel C shows the impact on firm performance for those firms whose CEO letters contained tech buzzwords. Post covid is a dummy variable equal to 1 for firms that publish CSR reports the period after the coronavirus pandemic (starting from the year 2020) and zero for the years prior. P ostCovid is a dummy variable equal to 1 for all firms post-pandemic, irrespective of whether they published a CSR report that year or not, and zero for the pre-pandemic years. The dependent variables are firm performance variables such as ROE, ROA, Tobin's Q, Net Profit Margin, and Operating Margin. Control variables include firm characteristics such as firm age, firm size, and firm liquidity. All regressions include firm, state, and year fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ROE	(2) ROA	(3) Tobin's Q	(4) Net Profit Margin	(5) Operating Margin
Panel A: Baseline					
post_covid	-0.030 (0.164)	0.003 (0.029)	0.965 (0.609)	-0.026 (0.054)	-0.003 (0.030)
Observations	500	501	500	501	501
Adjusted R ²	0.110	0.448	0.752	0.454	0.764
Panel B: Increase in Tech Positivity					
post-covid=1 × HighPosMedian=1	0.106 (0.079)	0.036*** (0.014)	-0.199 (0.293)	0.070*** (0.026)	0.041*** (0.014)
Observations	500	501	500	501	501
Adjusted R ²	0.112	0.456	0.752	0.463	0.768
Panel C: Increase in Uncertainty					
post-covid=1 × HighUncertMedian=1	0.095 (0.079)	0.033** (0.014)	-0.241 (0.295)	0.036 (0.026)	0.027* (0.014)
Observations	500	501	500	501	501
Adjusted R ²	0.115	0.457	0.751	0.457	0.771
Panel D: Increase in Strategy					
post-covid=1 × HighStrategyMedian=1	-0.085 (0.074)	-0.001 (0.013)	-0.576** (0.275)	-0.008 (0.024)	-0.023* (0.013)
Observations	500	501	500	501	501
Adjusted R ²	0.111	0.447	0.754	0.453	0.765
Panel E: Increase in Tech Uncertainty					
post-covid=1 × HighUncertTechMedian=1	0.148* (0.083)	0.020 (0.015)	-0.796*** (0.305)	0.058** (0.027)	0.013 (0.015)
Observations	500	501	500	501	501
Adjusted R ²	0.115	0.450	0.756	0.459	0.764
Panel F: Increase in Tech Strategy					
post-covid=1 × HighStrategyTechMedian=1	-0.073 (0.078)	0.003 (0.014)	0.141 (0.290)	0.022 (0.026)	-0.002 (0.014)
Observations	500	501	500	501	501
Adjusted R ²	0.110	0.447	0.752	0.454	0.763

Table 6: The post-2015 impact of COVID-19 on firm performance.

The table reports the impact of the COVID-19 pandemic on firm value and firm performance. Panel A shows the post-pandemic performance of all tech firms present in the S&P 500 index compared to pre-pandemic. Panel B reports the same but only for those firms that published ESG reports, while Panel C shows the impact on firm performance for those firms whose CEO letters contained tech buzzwords. Post covid is a dummy variable equal to 1 for firms that publish CSR reports the period after the coronavirus pandemic (starting from the year 2020) and zero for the years 2015-2019. PostCovid is a dummy variable equal to 1 for all firms post-pandemic, irrespective of whether they published a CSR report that year or not, and zero for the pre-pandemic years. The dependent variables are firm performance variables such as ROE, ROA, Tobin's Q, Net Profit Margin, and Operating Margin. Control variables include firm characteristics such as firm age, firm size, and firm liquidity. All regressions include firm, state, and year fixed effects. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ROE	(2) ROA	(3) Tobin's Q	(4) Net Profit Margin	(5) Operating Margin
Panel A: Baseline					
post-covid	-0.026 (0.113)	0.024 (0.017)	1.873*** (0.399)	0.002 (0.035)	0.012 (0.019)
Observations	430	431	430	431	431
Adjusted R ²	0.127	0.514	0.767	0.451	0.763
Panel B: Increase in Tech Positivity					
post-covid=1 × HighPosMedian=1	0.131 (0.088)	0.045*** (0.014)	-0.281 (0.313)	0.081*** (0.028)	0.048*** (0.015)
Observations	430	431	430	431	431
Adjusted R ²	0.130	0.528	0.766	0.463	0.769
Panel C: Increase in Uncertainty					
post-covid=1 × HighUncertMedian=1	0.088 (0.090)	0.031** (0.014)	-0.369 (0.322)	0.039 (0.029)	0.023 (0.015)
Observations	430	431	430	431	431
Adjusted R ²	0.133	0.520	0.766	0.453	0.774
Panel D: Increase in Strategy					
post-covid=1 × HighStrategyMedian=1	-0.093 (0.083)	0.006 (0.013)	-0.409 (0.293)	-0.002 (0.026)	-0.012 (0.014)
Observations	430	431	430	431	431
Adjusted R ²	0.127	0.513	0.767	0.450	0.763
Panel E: Increase in Tech Uncertainty					
post-covid=1 × HighUncertTechMedian=1	0.174* (0.092)	0.014 (0.014)	-0.819** (0.326)	0.046 (0.029)	0.002 (0.016)
Observations	430	431	430	431	431
Adjusted R ²	0.133	0.514	0.770	0.453	0.762
Panel F: Increase in Tech Strategy					
post-covid=1 × HighStrategyTechMedian=1	-0.083 (0.087)	0.012 (0.014)	0.116 (0.310)	0.032 (0.028)	0.011 (0.015)
Observations	430	431	430	431	431
Adjusted R ²	0.126	0.514	0.766	0.452	0.763

Online Appendix

Table A1: Technology Buzzwords.

The table shows the technology-related buzzwords used in our paper to signify futuristic technology, or the adoption of Internet 4.0 tools by technological companies. In our model, we included plurals, acronyms, and alternative words to capture sentiment for one particular technology. For example, Robotics includes robot, robots, and robotic, while Fintech includes financial technology and financial technologies.

Artificial Intelligence (or, AI)	Internet of Things (or, IoT)	Electric Vehicles (or, EV)	Data Analytics	Big Data
Machine Learning (or, ML)	Digital Transformation	Self-driving	Blockchain	

Table A2: Key variable definitions.

The definitions for five of our variables, extracted using Natural Language Processing (NLP) tools, is shown in the table below. We extract positive sentiments using the Roberta text classification tool. For the uncertainty and strategy sentiments, we first measure these sentiments overall across the entire document, and second, only for two sentences following and preceding tech-related words, thus capturing overall uncertainty and strategy, as well as tech-related uncertainty and strategy.

Variable	Variable Definition
Variable Definitions	
Tech Word Use	We measure the use of technology-related words by checking for the use of keywords first such as machine learning, artificial intelligence, the internet of things, digital transformation, digital, analytics, and electric vehicles. Second, we measured the sentiment around the use of tech words. Next, we calculate the number of tech words in a single CEO letter over the number of total words in the letter.
Tech Word Sentiment	To measure the sentiment around technology use, we measured the extent of positivity or negativity of two sentences before and after the use of a technology-related word. We measured this using a Roberta pre-trained model for sentiment classification. We focused on two sentences before and after the use of the keywords. The final technology-related positivity and negativity scores were calculated by averaging the score for all technology-related words found in the letter.
Uncertainty Score	To measure the uncertainty score, we counted the number of times the CEO letters had words related to uncertainty. The keywords were derived from Loughran and McDonald (2011). The final uncertainty score for each letter was taken by dividing the raw count of certainty words over the total number of words for each letter. In the data, we differentiate between the total uncertainty score of the entire letter, and the total uncertainty score of the sentences that appear before and after the use of technology-related words.
Strategy Score	Similar to the uncertainty score, we counted the number of times the CEO letters had words related to strategy. The keywords were derived from Loughran and McDonald (2011). The final strategy score for each letter was taken by dividing the raw count of strategy words over the total number of words for each letter. We used the same procedure from the uncertainty scores to differentiate between the total strategy score of the entire letter, and the total strategy score of the sentences that appear before and after the use of technology-related words.
World Events	We created a binary variable to indicate whether a CEO letter was written before (i.e., 0) or after (i.e., 1) COVID. In addition to this, we created a variable to count the extent to which a CEO letter mentioned world events such as Ukraine, Russia, or the COVID-19 pandemic.

Threat of AI – How Serious Can It Be?

Jinchang Wang

School of Business, Stockton University

Galloway, New Jersey 08215

jinchang.wang@stockton.edu

Abstract

ChatGPT and many other generative AI's have aroused hype of AI around the world. In this article, we put aside benefits that AI brings to us and focus on the potential threats that AI may pose to us. AI is threatening employment and social norms; it might be used by evil guys; and AI is surpassing humans in intelligence and may threaten the dominance position of humans on the earth. This paper evaluates how serious these threats can be. We argue that the impacts of AI currently on our society are controllable. And its threat to humanity in the future is not as serious as many people think since AI is not able to be self-conscious, therefore AI will remain as a tool of us and will not turn against us of its own accord.

Key words: Artificial intelligence, ChatGPT, Threat of AI, Self-Awareness, Humanity

ChatGPT's birth in 2022 reminded people that living with AI is not a dream in the distant future, instead, it is happening before our eyes now. We are witnessing an unprecedented event in the history of human beings: - For the first time, something is surpassing humans in intelligence. It has stirred hype in public media, predicting our life and the society at the time AI is around us in the future. Along with thrilling exhilarations, there are gloomy worries about the uncertainties of humanity. In this article, we put aside benefits that AI may bring to us and focus on the potential threats that AI may pose to us.

U. Is a Computer Going to Be Smarter than a Human?

Yes.

In fact, computers are already smarter than us in many aspects before ChatGPT, such as numerical computation in terms of speed and complication, memory capacity and memory accuracy, processing information from huge data sets as in big data and data mining. Deep Blue and AlphaZero beat human masters in chess. IBM Watson won Jeopardy!.

Computer intelligence is improving every day, while human intelligence remains unchanged. It seems a no-brainer to visualize the future when the computer surpasses human intellectually in all rounds.

II. Does Smart AI Pose a Threat to us?

Likely.

Jobs requiring intelligence tend to be replaced, maybe gradually, by smarter AI. It may not be too bad, but we'll lose the pride of handling challenging work as the most intelligent creature in the universe. Bad guys may use AI as a powerful destructive weapon. It would be even worse if AI is smarter than humans in all rounds. That is because less intelligent things tend to be dominated by those with higher intelligence, as chickens and dogs are controlled by humans. Confronting with AI, the best luck of less intelligent humans is being treated as pets, which is better than being slaved and killed.

III. Threats that AI Poses to Us

Threats of AI are in three aspects. First, AI poses a direct impact to our daily life. Second, AI is used as a powerful weapon by bad guys. Third, AI itself turns against us.

(3.1) AI threatens our jobs, privacy and security, and it will change our tranquil life.

Jobs requiring intelligence tend to be replaced by AI, such as play writing, translator/interpreter, receptionist, customer services, salesperson, professor, accountant, and secretaries.

In history, humans have survived such kind of threats from technology a few times. Examples include industrial revolution in which machines replaced handicrafts; nuclear energy development; using computers in numerical computation; using computers in text processing; and accomplishments of AI achieved before Open AI and ChatGPT such as Deep Blue, AlphaGo, and IBM Watson.

(3.2) “Smart AI” is used by evil guys as a powerful weapon.

AI can be used as a powerful weapon in physical battle ground as well as in computer systems, networks, and cyber world. AI weapons in five or ten years could be more devastating than any weapon that currently exists. In case such weapons were in the hand of evil guys, the consequence would be catastrophic.

AI is not the first technology that might be used by bad guys as a threatening weapon. Among “dangerous destructive armaments” are atom bombs, chemical weapons, bioweapons, cloning, stem cells, and brain waves.

There are always bad guys around. We cannot expect a world without evil. A necessary condition to reduce such threats is to curb bad guys' evil attempts. We have had experiences dealing with bad guys and protecting ourselves by regulations, rule of laws, police, and armed forces. We are not sure whether we could win in the next battle, but we are determined and capable to fight for win.

(3.3) AI turns against humans.

AI becomes dominant over humans on its own initiative. At that time, AI designs AI, AI produces AI, AI uses AI, AI controls AI, and AI determines its friends and enemies. As the result, we humans will face a “creature” that itself directly poses threats to us.

We have not experienced such threats in history. But it is plain to see the outcome of confrontation between two creatures of different intelligence levels – the dumber one is inevitably defeated. In front of a more intelligent creature, we’ll have no chance to win no matter how hard we struggle, just as chicken and pigs remain dominated by humans because of the difference in intelligence levels. Being dominated by a “machine” made by ourselves would be pathetic but seem inevitable.

IV. Will AI Replace Humans?

Man-made things have potentials to pose threats to humans *indirectly*, such as chemical products, explosives, nuclear weapons, stem cells, and AI. They are operated and controlled by humans. And humans are the “agents” using then or manipulating them. Strictly speaking, such a threat of AI is not from AI, but from humans. Dealing with such threats is an “internal” issue of our humans. We human beings have had experiences dealing with the threats from bad guys. We have laws, regulations, prisons, and armed forces to protect ourselves.

AI is different from the other man-made things. Robots may act by themselves. Robots may “think” by themselves. Once AI is smart enough to go by self, to research by self, and to replicate self, it is possible that it takes the initiative to turn against humans and dominate human beings. That would be a *direct* threat to us. Since AI is more intelligent, the situation would not be controlled by us. Human fate would be in the hands of AI. AI could put the human species to the brink of extinction.

AI is the first man-made technology that would threaten humans *directly by itself*, which causes a genuine threat. AI is likely the last invention of humans. Many scholars have warned of the gloomy fate of human beings confronting AI. For example:

Stephen Hawking:

“The development of full artificial intelligence could spell the end of the human race... It would take off on its own, and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn’t compete and would be superseded.” [5] “I fear that AI may replace humans altogether. If people design computer viruses, someone will design AI that replicates itself. This will be a new form of life that will outperform humans.” [4]

Nick Bostrom:

“If machine brains one day come to surpass human brains in general intelligence, then this new superintelligence could become very powerful. As the fate of the gorillas now depends more on us humans than on the gorillas themselves, so the fate of our species then would come to depend on the actions of the machine’s superintelligence.” [2]

James Barrat:

He called AI “our final invention”, and warned, “Artificial intelligence could drive mankind into extinction.” [1]

Will AI replace us as those scientists predicted?

AI is currently a smart but harmless “tool” of humans. Under what circumstance will AI turn against us?

V. Self-Awareness Is a Critical Intelligence

Anything ever created and made by humans is a “tool”, no matter how complicated it is. A tool is used by humans. A tool can be good or bad depending on the effect of using it, or on who uses it, a good guy or a bad guy. It will not go beyond human’s control as far as it remains a “tool”. Computers, explosive, and weapons are examples of such man-made tools.

A tool does not attack humans by itself because it is not self-aware. If a tool were self-aware, it would possibly tend to be disobedient and insubordinate. A necessary condition for computers to stop being a tool is: It becomes self-aware. A tool will never dominate its master as far as it is not self-aware. Self-consciousness would possibly “wake” a tool up so that it would be hard to reconcile itself to the role of a tool of complete obedience. Self-awareness gives rise to impetus of disobeying and dominating.

Self-awareness is “knowing existence of self subjectively”. It requires a higher level of intelligence. No computer so far has sense of “I”, so they remain as a tool like a hammer or a calculator, with no initiative or internal drive to “control” or “dominate” anything including humans.

VI. Will Computers Become Self-Conscious?

The answer to this question is: No. This is not a guess or a hypothesis. This is a proven Wang Theorem put forward by Wang [8]. The proof is concise:

Self-awareness is recognizing the existence of “self” which is a subjective consciousness. Everyone has a “self” which is unique. Since self-awareness is unique, it cannot be duplicated or copied. On the other hand, computer software is copiable by Church-Turing Theorem. So, copiable software cannot achieve uncopiable self-awareness. Therefore, self-awareness or self-consciousness cannot be realized by programming a computer.

Wang Theorem asserts that no matter how intelligent a computer is, it cannot be self-conscious. Self-awareness is a piece of human intelligence which cannot be achieved by any computer. Therefore, AI can never surpass human intelligence in all rounds. Human intelligence is mysteriously special because of self-consciousness. Higher intelligence level does not necessarily lead to self-awareness. Man-made intelligence is not logically possible to achieve self-awareness. Without self-awareness, AI, as man-made intelligence, would have no impetus to attack and dominate human beings.

VII. SCE, Self-Conscious-Emotions.

Some human consciousnesses are closely related to self-consciousness, which are known as *self-conscious-emotions*, *SCE*, in psychology [6]. Wang put human consciousness in three categories, strong-SCE, weak-SCE, and non-SCE depending on how closely it is related to self-awareness [7].

Examples of strong-SCE which is closely related to self-awareness:

Fear of death, free-will, sense of fairness, embarrassment, envy, guilt, shy, regret, reverence, shame, greed, pride, faith, dignity, sense of value of life, self-respect, self-restraining.

Fear of death is an essential strong SCE, which is an indication of our understanding of life and death. From this understanding, come the conceptions of soul, spirit, religion, sense of urgency, sense of inheritance, etc.

Examples of weak-SCE which is somewhat related to self-awareness:

Sincereness, sorrow, sorry, solitude, indignation, happiness, un-easiness, curiosity, sensitivity, hatred, obsession, responsibility, depression, humility.

Examples of non-SCE whose existence does not need self-awareness:

Hunger, thirst, pain, itching, fatigue, sleepy, joy, dread, orgasm, sexual desire, feeling warm and cold, tasting sweet, sour, salty, spicy.

Consciousnesses in non-SCE can be found in many animals. While strong-SCE only belongs to humans. Strong-SCE is highly dependent on self-consciousness. A creature that has no self-awareness does not have strong-SCE.

Logic does not preclude future AI from having “consciousnesses” of non-SCE. Logic precludes future AI from having SCEs especially strong-SCEs.

VIII. What Is AI without Self-Consciousness and SCE Like?

Let us imagine what will AI be like, which is super intelligence but is not self-aware:

AI is smarter and more knowledgeable than each of us, and smarter than all humans together, in all aspects. But it does not recognize itself.

AI does not fear death since it does not understand of live and death.

AI does not have free will.

AI had no soul and spirit.

AI has no internal impetus going to church.

AI is not self-motivated and self-spontaneous.

AI does not take the initiative to produce itself, so AI will not become a species that is able of self-replication.

AI is not self-thinking, self-acting, or self-restraining.

AI will not become a new species on the earth and in the universe.
AI will remain a tool of humans and never become humans.
AI will be “safe” to humans in the sense that it has no impetus to dominate the world and humans.

IX. Damages of AI without SCE Are Not as Serious

Something smarter than us does not necessarily pose a serious threat to us, as far as it remains a tool of us. We have already had computers “smarter” than us in numeric calculations and capacity of memory. AlphaGo already beat humans in all chesses. IBM Watson already won the knowledge game Jeopardy!. ChatGPT has shown the computers are more knowledgeable than any one of us. All these “smart things” are our tools. We will not lose control of them as far as they are not self-conscious. Since we have proved that AI will never be self-conscious, AI will remain a tool with no free will and SCEs and will not dominate human beings.

AI itself is not as threatening as many people think. Our attentions should focus on the people who use the tool of AI. The worry should be relieved that AI will enslave and dominate humans. People might be over apprehensive about our future.

X. Closing Remarks

ChatGPT is a milestone of AI, but it is not the first milestone of computer development. Prior to AI, we had milestones of computer’s numerical ability and milestone of computer’s text-processing ability, which marked progresses in phases of computer intelligence.

All the phasic improvements so far, including ChatGPT, are “quantitative improvements” which juxtapose or heap up technological capabilities in a computer to make its more efficient and handle more complicated scenarios. None of those quantitative improvements escalates a tool to a level of “non-tool” of self-spontaneousness. A tool does not pose serious threats to humans. It requires a “qualitative change” to fulfil the essential leap to make itself a genuine threat. Such qualitative change is “possessing self-consciousness”. However, as we have discussed in this article, self-consciousness is not achievable by AI.

2024 Nobel Physics laureate Geoffrey Hinton, known for his work in artificial neural networks and deep learning which earned him as “godfather of AI”, worries “we have not had such experience what will happen if something is smarter than we are. It may be wonderful. But there are many possible horrible outcomes, especially when losing control.” “It is still hard to tell how humans could prevent bad guys from doing bad things by using AI, which poses a genuine threat to human beings.” [3] We have showed that these threats are preventable and controllable compared to AI’s turning against us on its own initiative. Fortunately, AI can never be self-conscious, therefore AI will not have free will and SCEs, and will not attack humans of their own accord or on their own initiatives.

References:

- [1] [James Barrat 2013] Barrat, James, <Our Final Invention – Artificial Intelligence and the End of the Human Era>, Thomas Dunne Books, St. Martin's Press, 2013, New York
- [2] [Nick Bostrom 2014] Bostrom, Nick, <Superintelligence – Paths, Dangers, Strategies>, Oxford University Press, 2014.
- [3] [Hinton 2024] Hinton, Geoffrey, "Risks of Artificial Intelligence", Jan 16, 2024
<https://www.cbsnews.com/news/geoffrey-hinton-ai-dangers-60-minutes-transcript/>
- [4] [Stephen Hawking 2017] "An Interview with Wired", Wired, Nov. 2017
- [5] [Stephen Hawking 2014] "An Interview with the BBC", Dec. 2014
- [6] [Tracy and Robins 2004] Tracy, J. and R. Robins, Putting the Self into Self-Conscious Emotions: A theoretical Model, Psychological Inquiry, Vol. 15, No.2, p.103-125, 2004
- [7] [Wang 2018] Wang, Jinchang, Unachievable Consciousness on Electronic Robots, Proceeding of Northeast Decision Science Institute NEDSI, Providence, Rhode Island, April 2018
- [8] [Wang 2013] Wang, Jinchang, On the Limit of Machine Intelligence, International Journal of Intelligence Science, Vol. 3, p.17

Ethics of AI in Education

Alexander Pelaez – *Hofstra University – Hempstead NY 11549* – Alexander.Pelaez@hofstra.edu

Amal Jacobsen – *Progressive School of Long Island- Merrick, NY 11566*, amal@progressiveli.org

Jim Samuel – *Rutgers University, New Brunswick NJ 08901*, jim.samuel@rutgers.edu

Kaushik Sengupta – *Hofstra University – Hempstead NY 11549* – Kaushik.Sengupta@hofstra.edu

Introduction

As professionals and consumers embrace AI technologies, academic institutions and faculty struggle to find balance between the constructive use of AI as a tool for engaging students and limiting AI to prevent academic integrity issues. Students face serious consequences if they are caught using a technology whose use has become endemic on college campuses; however, there is still a tendency for students to use the technology to achieve an edge. While academic institutions have appealed to the ethical consciousness of their students for generations to not engage in academic fraud, suddenly those same institutions are finding their ethical appeals falling on deaf ears in a way perhaps never seen before. These challenges demonstrate the crucial need to re-evaluate the ethical landscape that has been upended by the advent of AI chatbots.

While cheating, fraudulence and plagiarism have always plagued college campuses, the challenges presented by the use of AI chatbots is perhaps unprecedented. This crisis forces us to ask what, exactly, has changed, besides the advent of a new technology. Has the ease of cheating increased such that it has naturally become a more widespread problem, or has the pervasiveness of AI chatbots fundamentally shifted the ethical demands being set upon students? Have students simply become more dishonest, or does the crisis created by the widespread use of AI chatbots reveal a deeper issue at the rudiment of the educational system?

This paper attempts to identify the source of these issues by proposing several research questions through the lens of three different ethical approaches, consequentialism, deontology, and relativism. In exploring these questions, the aim is to shed light upon the core of students' and educators' competing priorities and beliefs about this emerging technology and its disruptive capacity.

Literature Review

The entire field of education is being shaken by the advent of AI chatbots and their widespread use in the classroom (Dimitriadou and Lanitis 2023). Educators and students are using these technologies in a variety of ways and its proliferation is seen as problematic by some. According to the Pew Research Center (2024), one-third of high school teachers responded that the AI tools are more harmful to education, and Inside Higher Ed (2023) reports almost half of college faculty report negative effects of AI on education. The balance between allowing students to use technologies that will most certainly become part of their occupation and promoting academic integrity becomes the focal point for educators, who strive to impart knowledge and critical thinking to their students.

Pew Research Center (2023) found that two-thirds of high school students are aware of technologies such as ChatGPT, and their attitudes towards its usage is mixed. While 69% of those surveyed found it acceptable to use to research new topics, 57% found it unacceptable to use to write essays. Intelligent.com (2023) similarly found that 69% of students used it for writing assignments while 74% consider its use as some form of cheating. The statistics show the ethical dilemma students face with respect to the use of these tools.

Whilst a century ago the university was an elitist conclave focusing on liberal arts education, with its focus on creating “well-rounded individuals”, today students see higher education as stepping stones on the path towards a financially remunerative career. Degrees, rather than being signifiers of the richness of a well-rounded education, are job market necessities in the information economy. Students are trained from before the time of high school to see higher education and its degrees in this way (Brooks et al., 2021). For this reason, education has ceased to be seen as an end unto itself, but a means to an end. Crucially, these degrees are not seen as practical necessities because of the knowledge gained in such educational programs, but because the degrees themselves are the keys to unlocking career paths desired by the

students (Brooks et al., 2021). The actual educational lessons are valuable only insofar as they lead to the awarding of a degree, and are not valuable in-and-of themselves. Such a situation, even before the advent of AI, was a powder keg, ready to explode.

With the advent of AI technology, students are put into a situation of conflicting ethical demands. The university, professors, and teachers are tending to make deontological claims to appeal to the ethical consciousness of their students, i.e. cheating is wrong, therefore you shouldn't do it. The students, meanwhile, often see AI technology in a utilitarian way. By using AI technology, they can more easily obtain the degree they are seeking. If degrees are more valued than the actual education they are receiving, why wouldn't a student want to use every tool at their disposal to attain what they need to succeed? This kind of cynically utilitarian ethos might offend an educator, but students have been trained for decades to see education in exactly this light.

Ethics based theories

The ethics literature is rich with a variety of ethical positions, however we focus on three primary ethical theories: consequentialism, deontology, and relativism. Our aim is to assess AI shapes student's decision making, and what perspectives they have of the interaction for given outcomes.

Consequentialism / Utilitarianism

Business practices and actions are generally consequence driven (Gustafson, 2018) Utilitarianism as a form of consequentialist ethics, judges actions by whether they contribute the maximum amount of good. Many students are able to cynically place the advancement of their career as the "good" they are aiming at, without much ethical reflection (Peterson, 2003; Gustafson, 2018). In a consequentialist ethos, actions are morally judged according to their outcomes (Peterson, 2003) . For students who are trained from an early age to enter into an outcome-driven world, this is often a natural

fit. Cheating is wrong insofar as being caught will harm your ability to thrive in an academic environment.

Even from a strictly consequentialist point of view, however, major questions arise as to the academic outcomes of an environment dominated by the use of AI chatbots to grease the wheels of academic striving. As students enter into a world in which both employees and employers, both companies and consumers, are increasingly relying on AI to cut down on workloads, the question arises as to whether the widespread use of AI is in fact improving outcomes for all parties involved, or whether such an environment just degrades or stifles overall standards of excellence, nuance, creativity and critical thinking. According to a report by Europol, by 2026 as much as 90% of all online content will be AI-generated. If we are to judge actions purely by their outcomes, the question arises as to whether this is in fact the most desirable outcome.

Deontological

As previously cited, 74% of surveyed students view the use of AI chatbots as some form of cheating. This shows that in the back of most students' minds, consequentialist rationalizations cover up a deontological awareness that the ethics of using AI chatbots is murky at best. Deontology, as famously summarized by Immanuel Kant's categorical imperatives, is the idea that something is morally right only if we could envision everyone else acting in a similar way (Benabou et al., 2024; Megias et al. 2023). If everyone cheated, that would be bad, so cheating is universally wrong from a deontological point of view. This point of view would suggest that regardless of outcomes, people should guide their behavior by a sense of ethical duty.

Despite the consequences that professors and academic institutions threaten their students with, educators are largely relying on the deontological conscience of their students when they try to instill the sense that the use of AI chatbots is cheating, and should therefore be avoided. The problem is that because students have competing

ethical motivations, deontological imperatives can be seen by ambitiously striving students as toothless and naive.

Relativism

Despite how much students will often internally rationalize their behavior in consequentialist terms, most students are actually using a relativistic ethos to frame their choices. Relativism is the belief that actions should be judged according to the environment and social situations, and not in absolute terms (Shahryari, 2023) . Very simply, the widespread justification that “everybody else is doing it.” This is perhaps one of the greatest single difficulties that educators face. When the use of AI chatbots is so pervasive, so omni-present in the educational and business world alike, it becomes exceedingly difficult to appeal to any other higher ethical principles. By not participating in a culture that everybody else plainly is, students easily see that they are, very simply, putting themselves at a disadvantage by not using the tools that are so widely available and readily used by others around them. In fact, many students might openly criticize professors or academic institutions for not preparing them for the world as they will actually encounter it after graduation, since the use of AI chatbots is so common. If educators wish to have any impact upon the ethical consciousness of their students regarding the use of AI chatbots, therefore, it is perhaps of the highest priority to address this relativistic ethos first and foremost. To not do so would be to ignore the basic reality their students are operating within.

Research Questions:

The proposed research aims to develop a survey instrument along with a proposed assignment or case study to students to gauge how students perceive situations. From the literature review we therefore create the following research questions:

First, we identify the nature by which students believe their assignments impact their long term outcomes and then attempt to align them to a set of consequences, perceived

or real. Students who believe in a form of an “end justifies the means” approach will most likely align to a consequentialist view of the use of AI. Therefore, we develop the following research question:

RQ1: *What level of consequences appear to impact the students’ use of AI?*

RQ2: *What are the students’ perceptions of long term consequences , such as employment risks, by using AI in the classroom?*

Next, we aim to explore the students’ deontological view of the use of AI in the classroom. Are students bound by a set of morals that prevent them from using AI because they break rules, either explicit or implicit. We can assess whether students take an individualistic view or collective view of the ethical dilemma, i.e. do they feel any responsibility if others do not use AI. We can then state the following research questions:

RQ3: *Is there a moral question using ChatGPT for one’s own academic learning?*

RQ4: *Are there integrity issues using ChatGPT even if it is allowed?*

Finally, we focus on relativism, in which one’s own beliefs are based on the situation and environment they are in. We aim to identify situations in which a relativist view may impact a student’s use of ChatGPT. For example, when faculty may be remiss in identifying the use of chatgpt, either by choice or (need a better word for not by choice), and students are widely using ChatGPT, does this affect one’s own decision. Therefore, we ask the following:

RQ5: *Do other students’ actions affect one’s own belief in the use of ChatGPT?*

RQ6: *Does an educator’s attitudes towards ChatGPT affect one’s own belief in the use of ChatGPT?*

Methodology

Based on preliminary research, we aim to provide a sample assignment that uses ChatGPT currently in development. Using this as a foundation, we will develop a questionnaire to assess the student's ethical perceptions of the use of AI technologies in the classroom. We use a ChatGPT enabled set of assignments across multiple courses where students are asked to get their initial responses on the assignment using ChatGPT and then use a rubric to critique the quality of the responses from ChatGPT and suggest improvements in a multi-step process (Sengupta and Pelaez, 2024). The suggested improvements must map back to the course's concepts and discussions. Evaluation of the students' quality of output on this assignment is based on a qualitative evaluation by the instructor on a rubric that is shared with the students. The assumption here is based on the premise that students will actually use ChatGPT responses to critically analyze these responses on their own based on their learning from the course and improve their knowledge about the concepts discussed in the course. The pool of students will span the domain of university students including graduate and undergraduate across disciplines such as management and business analytics. Furthermore, we aim to explore the attitudes of students across a spectrum of grade levels and capabilities, including high school students, who are increasingly leveraging these tools, altering the questions and the assignment to be more age appropriate, but providing us with insight into student's perceptions at multiple levels.

Discussion

By understanding the ethical perceptions of students, educators and administrators can be more acutely aware of the impact on a student's educational outcomes. This will enable educators to develop better strategies, assignments and other assessments towards the ultimate goal of ensuring proper outcomes, but also serve as a means to leverage AI tools that will be prevalent in the course of a student's career.

References

- Bénabou, R., Falk, A., & Henkel, L. (2024). *Ends versus means: Kantians, utilitarians, and moral decisions* (No. w32073). National Bureau of Economic Research.
- Brooks, R., Gupta, A., Jayadeva, S., & Abrahams, J. (2021). Students' views about the purpose of higher education: a comparative analysis of six European countries. *Higher Education Research & Development*, 40(7), 1375-1388.
- Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, 10(1), 12.
- Gustafson, A. (2018). Consequentialism and non-consequentialism. In *The Routledge companion to business ethics* (pp. 79-95). Routledge.
- Megías, A., de Sousa, L., & Jiménez-Sánchez, F. (2023). Deontological and consequentialist ethics and attitudes towards corruption: a survey data analysis. *Social Indicators Research*, 170(2), 507-541.
- Inside Higher Ed. (2023, October 31). *Most students are outrunning faculty on AI use*. Retrieved January 12, 2025, from <https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2023/10/31/most-students-outrunning-faculty-ai-use#:~:text=Attitudes%20and%20Policies,varied%20on%20implementing%20systemwide%20policies>.
- Intelligent.com. (2023). *4 in 10 college students are using ChatGPT on assignments*. Retrieved January 12, 2025, from <https://www.intelligent.com/4-in-10-college-students-are-using-chatgpt-on-assignments/#:~:text=ChatGPT%20most%20commonly%20used%20for,know%20what%20the%20tool%20is>.
- Peterson, M. (2003). From consequentialism to utilitarianism. *The Journal of philosophy*, 100(8), 403-415.
- Pew Research Center. (2024, May 15). *A quarter of U.S. teachers say AI tools do more harm than good in K-12 education*. Retrieved January 12, 2025, from <https://www.pewresearch.org/short-reads/2024/05/15/a-quarter-of-u-s-teachers-say-ai-tools-do-more-harm-than-good-in-k-12-education/>
- Pew Research Center. (2023, November 16). *About 1 in 5 U.S. teens who've heard of ChatGPT have used it for schoolwork*. Retrieved January 12, 2025, from <https://www.pewresearch.org/short-reads/2023/11/16/about-1-in-5-us-teens-whove-heard-of-chatgpt-have-used-it-for-schoolwork/>
- Sengupta, K., & Pelaez, A. (2024). *AI Tools as a Teaching Aid for Graduate Level Courses in the MBA Curriculum*. 2024 Annual Conference of the Decision Sciences Institute. <https://decisionsciences.org/>
- Shahryari, S. (2023). Absolutism, Relativism, and Pragmatic Fallibilism: A Reply to Stump. *Journal for General Philosophy of Science*, 54(2), 331-338.

DataSciLawBot:
An LLM Chatbot for the Application of Data Science Industry Laws to Business Decisions

Sawyer Huang
New York University
gh2348@nyu.edu

Andres Fortino
agfortino@nyu.edu
New York University

Roy Lowrance
New York University
roy.lowrance@gmail.com

ABSTRACT

This study addresses the challenge of optimizing data science project management through the application of data science laws and principles. The proposed solution, DataSciLawBot, leverages a Language Model (LLM) technology to provide immediate, actionable guidance, enhancing the decision-making process in data science projects. The proof of concept aimed to demonstrate the feasibility and effectiveness of the DataSciLawBot in a real-world setting. In a trial involving case studies and expert comparisons, the DataSciLawBot was assessed against standard practices for managing data science projects. The results suggest that less experienced practitioners can improve their efficiency in discovering and applying relevant laws to data science project management tasks. Our conclusions were reinforced by the strong alignment between the chatbot's recommendations and those of subject matter experts, as measured by Cohen's kappa. The main conclusion is that DataSciLawBot effectively bridges the gap between data science best practices application and data science practical project management, leading to more streamlined processes. The implications of these results suggest substantial potential for enhancing data science project management frameworks within the data science community. Future work could expand the chatbot's capabilities to cover more complex scenarios and broader applications, potentially across various industries. The main contribution of this project is to show how an LLM can be used to improve the management of data science projects, suggesting that a similar approach would be beneficial for other kinds of projects requiring extensive domain knowledge.

INTRODUCTION

Data science projects often grapple with inefficiencies and regulatory challenges that can hinder their progress and impact overall performance. A significant obstacle is a need for data science project managers to navigate the extensive array of data science best practices and practical guidelines often codified into what might be considered industry laws. These "laws," or best practices, are not only complex but also continually evolving. This situation presents an opportunity to develop a tool that can provide real-time, accurate, and actionable guidance to improve project outcomes by applying these practical laws.

Introducing DataSciLawBot as a solution to these challenges. This chatbot utilizes a specialized LLM to assist data science team managers by offering guidance based on the "laws of the data

science industry.” These laws are designed to improve not only technical execution but also enhance overall project success through better management practices. This approach has been successfully implemented in other technical fields, particularly in information technology (Fortino, 2024) and information security (Yu, 2024). The goal of this project was to replicate that success within the data science discipline.

DataSciLawBot utilizes state-of-the-art LLM technology, capable of understanding and applying a curated database of data science laws and best practices. This technology allows the chatbot in turn to analyze project-specific situations and provide tailored advice that aligns with established guidelines and best practices, thus ensuring that data projects are not only compliant but also optimally managed.

The deployment of DataSciLawBot offers several key benefits: (1) Enhanced Project Efficiency: Automating the application of data science laws significantly reduces the time data science project managers spend on regulatory compliance; and (2) Improved Decision-Making: With instant access to relevant guidelines, project managers can make informed decisions that align with best practices, reducing the risk of costly mistakes.

This project explores whether an LLM-based chatbot can effectively streamline the management of data science projects by providing automated, real-time guidance on applicable laws and best practices. The research focuses on the chatbot’s ability to improve data science project management outcomes and compliance with industry standards.

This project’s primary contribution lies in its innovative use of LLM technology to improve data science project management. By automating the application of complex laws and guidelines, DataSciLawBot helps streamline project workflows, marking a significant advancement in how data science projects are managed.

We accomplished this in four steps: (1) Create a set of functional requirement specifications for the DataSciLawBot project vetted by the SME; (2) Develop a comprehensive database of data science laws and principles, also vetted by the SME; (3) Create the DataSciLawBot chatbot; and (4) Validate the chatbot’s effectiveness, usability, and accuracy using human expert baseline.

The completion of this project is of significant importance to STEM educational efforts as well. The project aims to develop DataSciLawBot, a tool designed to enhance the management of data science projects by applying specialized knowledge in data science laws and best practices. For a data science faculty member, this tool represents an innovative blend of technology and education, reinforcing their mission to prepare students for the complexities of modern industries through hands-on experience with cutting-edge technologies. For example, DataSciLawBot can help faculty design case studies that reflect real-world challenges in data science project management, allowing students to practice applying legal and best-practice principles in a simulated environment. Additionally, the bot can be integrated into coursework to provide instant, interactive feedback on student projects, fostering deeper understanding and engagement with industry-relevant scenarios. The development and implementation of DataSciLawBot align with the vision of integrating technology into business practices, providing a practical tool that can be used within the university curriculum.

The goal of DataSciLawBot: An LLM Chatbot for Data Science Law Application was to develop a specialized chatbot that utilizes Language Model (LLM) technology to simplify the process of

applying data science laws and principles and thereby re-risking the discovery of appropriate industry laws. By providing data scientists with instant access to relevant laws, DataSciLawBot aims to enhance the efficiency and compliance of data science project management. We also found that in some cases, the LLM found applicable laws that the human expert had overlooked.

Research Questions

In this paper, we will examine two research questions comparing an AI model's ability to select applicable data science laws versus human expert performance in managing data science projects.

Research Question 1: How feasible is it to build a generative AI-powered chatbot that can accurately deliver context-specific guidance on data science project best practices to professionals in real-time?

This question aims to evaluate the technical feasibility of developing the chatbot. It involves assessing whether the chatbot can be built to provide precise and relevant advice by comparing its responses with those of subject matter experts (SMEs).

Research Question 2: How well will such a chatbot perform compared to human SMEs?

This question explores the validation process of the chatbot's feasibility and effectiveness. It involves using Cohen's Kappa metric to measure the level of agreement between the chatbot's responses and those of subject matter experts, ensuring that the chatbot's outputs are consistent with expert judgment. A confusion matrix is used to analyze the correctness of the chatbot's predictions, providing a detailed view of its accuracy in delivering relevant advice when compared to the analysis of an SME. By leveraging these validation metrics, the research seeks to confirm that the chatbot can reliably and accurately assist information security professionals.

Hypothesis

The integration of advanced technologies in data science management is critical for enhancing the efficiency and accuracy of professional practices. This study investigates whether building a generative AI-powered chatbot is feasible and if it can significantly improve decision-making processes and knowledge management within the field.

H1: An LLM-based chatbot fine-tuned with data science laws will apply these laws to solve data science cases just as well as a human subject matter expert.

This hypothesis posits that the use of a generative AI-powered chatbot will demonstrate feasibility by showing substantial agreement with subject matter experts' answers, as validated by Cohen's Kappa metric and the confusion matrix. The validity of this hypothesis will be tested through structured technology trials, comparing the chatbot's performance against traditional methods of information retrieval and application in the information security domain. The test is the solving of use case scenarios where data science management principles need to be brought to bear to solve the problem, such as in a Harvard case study, for example.

LITERATURE REVIEW

The primary objective of the review was to establish a comprehensive understanding of the current state of AI chatbot technology and its applications, providing a solid foundation for the proposed implementation of DataSciLawBot. We examined studies and use cases where similar technologies have been applied, assessing their outcomes, and drawing parallels to the envisioned application in the data science domain. The scope of this review extends to covering pertinent aspects such as the evolution and capabilities of AI chatbots, their successes and limitations in different industries, and the potential for adaptation and innovation within the context of data science project management.

The data science industry is rapidly expanding, with a significant increase in demand for skilled professionals capable of extracting insights from large datasets. According to the U.S. Bureau of Labor Statistics, the field of data science is projected to grow by 15% from 2019 to 2029, which is much faster than the average for all occupations (US-BLS, 2019). This growth is fueled by the rising need for data-driven decision-making in various sectors, including finance, healthcare, and technology. Data science integrates complex algorithms, advanced analytics, and machine learning techniques to tackle large-scale data processing challenges, making it an invaluable asset for organizations aiming to leverage big data for strategic advantages.

Within this industry, project management plays a crucial role in guiding data science initiatives towards successful outcomes. The Project Management Institute (PMI) highlights that effective project management in data science involves navigating unique challenges such as handling unstructured data, ensuring data security and privacy, and managing interdisciplinary teams with varied technical expertise. Despite the growing adoption of project management practices, many data science projects still face high rates of delays and failures. A study by Gartner estimated that through 2022, 85% of AI projects would deliver erroneous outcomes due to bias in data, algorithms, or the teams managing them. This statistic underscores the importance of robust project management in mitigating risks and guiding data science projects to fruition.

The application of AI and machine learning in project management is emerging as a game-changer, offering solutions to some of these challenges. AI-driven tools like chatbots are increasingly being deployed to streamline project management processes, enhance team collaboration, and provide real-time decision support. Luo et al. (2022) underscore the diverse applications of chatbots across business sectors, suggesting their potential in enhancing business processes, including project management. Furthermore, Santos et al. (2022) illustrate the successful application of a conversation-driven approach in chatbot management, improving user interaction and engagement in project settings. Additionally, a study by Hassani & Silva (2023) emphasizes the revolutionary role of conversational AI tools like ChatGPT in transforming data science fields, providing further validation for the integration of AI in data science project management. These developments point towards a growing trend of integrating AI tools like chatbots into project management frameworks to enhance efficiency and decision-making in data science projects.

The integration of advanced AI capabilities in project management, particularly in the data science domain, is redefining industry standards, marking a significant shift from traditional management methodologies to more agile, AI-assisted approaches.

In the realm of data science, a significant problem lies in the efficient management of projects, particularly in terms of selecting appropriate methodologies and adhering to evolving data

science laws and principles. Despite the advancements in technology and data processing capabilities, data scientists often grapple with the complexities of project management, including staying abreast of legal regulations and systematic methodologies. This challenge is not merely theoretical; a Gartner study revealed that through 2022, around 85% of AI projects would produce erroneous results due to biases in data or methodologies, highlighting the need for more structured and informed approaches in data science project management. Furthermore, Ferreira et al. (2023) emphasize the disconnect in agile project environments, particularly in data science and process automation, where the lack of direct communication with end-users often leads to misalignment with actual project needs and requirements. This situation underscores the need for a solution that bridges the knowledge gap and streamlines project management processes. The problem, therefore, is not just managing the vast and complex data but doing so in a way that is legally compliant, methodologically sound, and aligned with the end-users' needs, making it a crucial area for improvement in the field of data science.

The proposed solution to the problem of inefficient data science project management is the implementation of an AI-driven chatbot called DataSciLawBot. This chatbot is designed to provide real-time guidance and support in selecting appropriate methodologies and adhering to data science laws and principles. The solution leverages advancements in artificial intelligence, particularly natural language processing and machine learning, to offer a conversational interface that assists data scientists in navigating the complexities of project management. Unlike traditional project management tools, which often lack real-time interactivity and best practice insights, DataSciLawBot aims to offer a more dynamic and informed approach.

In the realm of AI and chatbots, several studies provide evidence of the successful application of similar technologies in addressing complex problems. Luo et al. (2022) reviewed various state-of-the-art chatbot designs and their applications across different business sectors, demonstrating how chatbots have significantly enhanced business processes and decision-making. Their findings underscore the adaptability and effectiveness of chatbots in diverse industry contexts. Similarly, Santos et al. (2022) presented a conversation-driven approach for chatbot management, showing how strategic content management and user interaction analysis led to improved engagement and effectiveness. Extending this further, a study by Wilkinson (2024) discussed the integration of AI chatbots like ChatGPT in education, emphasizing their potential to enhance learning experiences and facilitate complex tasks, offering parallels for DataSciLawBot in data science project management. These examples indicate that AI-driven chatbots can effectively manage and streamline complex processes, making them well-suited for tackling the challenges in data science project management.

By integrating an AI-driven chatbot into the data science workflow, the DataSciLawBot project seeks to offer a novel solution that enhances efficiency, ensures compliance, and aligns with user needs, thereby addressing the gaps identified in current project management practices.

The core technology proposed for the DataSciLawBot project is an advanced Generative AI-driven chatbot utilizing inherent natural language processing (NLP) and machine learning capabilities of Generative AI. This technology stands out for its ability to process and interpret human language, allowing for interactive, conversational engagement with users. NLP enables the chatbot to understand queries in natural language, analyze them, and provide relevant, context-specific responses, making it a potent tool for real-time assistance in project management. Machine learning, particularly reinforcement learning as discussed by Kintsakis, Psomopoulos, and Mitkas (2019), underpins the chatbot's ability to learn from interactions, continuously improving its response accuracy and relevance. This adaptive learning capability

ensures that the chatbot evolves with the project's needs, offering tailored guidance and support. The integration of these technologies into a chatbot like DataSciLawBot is novel in the context of data science project management, but their use in facilitating complex tasks has been demonstrated in various fields, including healthcare and e-commerce, as shown in studies by Fan et al. (2021) and Hsu et al. (2023).

Additionally, the relevance of Generative AI chatbots in enhancing communication and managing workflows has been explored in the healthcare sector, as evidenced by the work of Battineni, Chintalapudi, and Amenta (2020). Their study on the design of AI chatbots during the coronavirus epidemic demonstrates the versatility and impact of chatbots in critical situations, offering insights into their potential applications in other complex domains like data science. The technology's application in DataSciLawBot is expected to revolutionize data science project management by providing a tool that not only assists with project-specific queries but also ensures compliance with relevant laws and principles. The versatility of AI chatbots in various industries, as reviewed by Luo et al. (2022), indicates the maturity and effectiveness of this technology in managing complex information and decision-making processes. This maturity suggests that implementing such a system in the data science domain is feasible and likely to yield positive results. The success of similar AI applications in different settings like the Evataalk chatbot for educational purposes, described by Santos et al. (2022), provides a solid foundation for anticipating the successful adaptation and implementation of this technology in data science project management, ensuring that DataSciLawBot is not only a viable solution but also an innovative approach to tackling the unique challenges of this field.

Use Cases

The implementation of AI-driven chatbots, akin to the proposed DataSciLawBot, has been successfully realized in various industries, offering compelling use cases that reinforce the viability and effectiveness of this technology. Two notable examples in the fields of education and healthcare demonstrate how chatbots have been employed to address complex challenges.

Use Case 1: Education In the education sector, AI chatbots have been utilized to enhance learning experiences and assist both students and educators. Chen et al. (2023) conducted a study on the application of AI student assistants in classrooms, illustrating how chatbots can support student success. Their research involved developing a chatbot that helped students learn basic AI concepts. The chatbot proved to be an engaging and responsive tool, significantly aiding in the learning process. The success of this implementation in an educational setting, where the chatbot facilitated interactive learning and provided accessible educational resources, aligns closely with the goals of DataSciLawBot in data science project management. This use case exemplifies how chatbots can effectively disseminate complex information and interactively engage with users, ensuring knowledge transfer and support in data-driven environments.

Use Case 2: Healthcare In healthcare, chatbots have been effectively used for patient engagement and support, particularly in remote or resource-limited settings. Fan et al. (2021) explored the use of a self-diagnosis health chatbot in real-world scenarios. Their case study showed that the chatbot was instrumental in providing medical guidance and support, handling a wide range of medical queries. The chatbot's ability to offer immediate, confidential, and accurate responses to health concerns mirrors the objectives of DataSciLawBot in providing real-time, reliable assistance in data science projects. The healthcare chatbot case study underscores the potential of chatbots in managing sensitive and complex information, offering a

relevant parallel to the application of DataSciLawBot in navigating the intricate landscape of data science laws and methodologies.

These use cases demonstrate the successful application of AI chatbots in fields that require precise information handling, responsive interaction, and adaptive learning capabilities. The experiences from these implementations provide a solid foundation for the anticipated success of DataSciLawBot in the domain of data science project management.

From these examples and existing literature supporting this application, we conclude that the integration of AI-driven chatbots, particularly in domains requiring complex decision-making and knowledge processing, holds significant potential in enhancing efficiency and accuracy. The adaptability of chatbots, as evidenced by Luo et al. (2022), and their successful applications in diverse fields, ranging from education to healthcare, as seen in the studies by Chen et al. (2023) and Fan et al. (2021), underline the broad applicability and effectiveness of this technology. Specifically examining the intersection of AI chatbots with data science project management, a field characterized by its requirement for precision, compliance, and rapid adaptation to evolving data laws and principles, reveals a promising avenue for the implementation of our proposed DataSciLawBot.

Most of the literature reviewed concluded that the use of AI chatbots not only enhances user interaction and engagement, as demonstrated by Santos et al. (2022), but also facilitates the management of complex tasks, thereby improving overall workflow efficiency. This is particularly relevant for the DataSciLawBot project, as the reviewed studies provide insights into the potential of AI chatbots to streamline data science project management and providing systematic methodological guidance. The successful use cases in various sectors reinforce the viability of applying a similar AI chatbot technology in the data science domain, suggesting that the proposed DataSciLawBot has the potential to address the specific challenges identified in this field effectively.

In conclusion, the result of this review for the DataSciLawBot project is promising. It validates the choice of technology and approach, aligning with successful implementations in other domains while also highlighting unique considerations for its application in data science project management. The synthesis of these findings forms a compelling argument for the feasibility and potential success of the DataSciLawBot, setting a foundation for the next steps in its development and implementation.

APPROACH AND METHODOLOGY

Data science projects frequently encounter challenges related to efficiency and compliance due to the complex nature of data science laws and best practices. There is a significant need for tools that can seamlessly integrate these laws into daily project management tasks to enhance project outcomes.

Problem Statement and Research Question

The project addresses the need for data science professionals to efficiently apply industry best practices in their technology projects. The solution involves the development of an advanced chatbot using large language models (LLMs). This tool is designed to assist professionals by making information about these best practices readily accessible, accurate, and comprehensive, thereby enhancing compliance and decision-making in technology adoption.

The proof of concept for the chatbot focuses on validating its ability to integrate and utilize a comprehensive database of best practices in data science. Initially, the GPT model was implemented using only the collected database of 100 data science best practices (referred to as “laws”). However, the results were not satisfactory; the GPT model attempted to find laws by itself, leading to a significant gap between the GPT’s answers and those of subject matter experts (SMEs).

The first step was to validate the 100 data science laws with subject matter experts to determine if all the “laws” were directly relevant to data science project management. Some laws inform building machine learning models, for example, but not useful in managing data projects. Others were more applicable to other sectors of technology, such as information security. The reduced set of applicable laws we developed was 54 in number and presented in Appendix B.

The selection of the subject matter expert (SME), who is a former director of a Data Science graduate program at a major university and a professor of data science at a prominent research university, was crucial in ensuring the expertise and relevance of feedback in refining the chatbot’s responses. We are planning rigorous testing phases that will focus on the chatbot’s performance in terms of accuracy and response quality. This testing phase will involve real-world scenarios to ensure the chatbot meets the functional requirements and delivers high-quality answers. With the help of the SME five Harvard Business Review case studies presenting complex data science project management scenarios and issues to be resolved were selected.

Database of Industry Laws

The initial phase of this research was a comprehensive review of the expansive realm of potentially usable data science laws and best practices. Given the voluminous nature of information in the data science domain, as highlighted in the introduction, this review was critical to identifying the most relevant and impactful practices that could aid professionals in their daily security management tasks. The challenge was not only to collect these practices but also to ensure they were up-to-date and adaptable to the rapid changes in the data science landscape.

The Need for a Taxonomy

The initial implementation of the DataSciLawBot chatbot using the ungrouped database of 54 data science laws and principles posed significant challenges. When a subject matter expert (SME) and the chatbot were tasked with selecting the top five laws applicable to a specific case, there was a notable discrepancy between their choices. The low agreement between the SME and the chatbot highlighted the difficulty in effectively navigating and applying the extensive collection of laws without a structured framework. This inconsistency in law selection indicated a need for a more organized approach to assist both human experts and the chatbot in identifying the most relevant principles for a given scenario.

Developing the Taxonomy

A taxonomy was developed to organize the 54 data science laws and principles into distinct clusters to address the mismatch between the SME and the chatbot. The process involved carefully analyzing each law and principle, identifying common themes and objectives, and grouping them accordingly using a prompted LLM (A typical prompt for this purpose is given in

Appendix A.) The resulting taxonomy consisted of nine clusters: Model Selection and Optimization, Data Quality and Preprocessing, Statistical Principles and Interpretation, Causal Analysis and Relationships, Human Judgment and Cognitive Biases, Data Analysis and Visualization, Data Engineering and Architecture, Agile Development and Technical Debt, Uncertainty and Complexity Management. This structured approach aimed to provide a decision tree that could guide users in selecting the most appropriate cluster(s) based on the specific case at hand and then focusing on the relevant laws within those clusters to identify the top five most applicable principles. The resulting taxonomy is provided in Appendix C.

Improved Chatbot and SME Matching with Taxonomy

After implementing the taxonomy, both the subject matter expert (SME) and the chatbot were given access to this structured framework along with the individual laws. The taxonomy acted as a guide, allowing users to first identify the general area of laws applicable to a given case by selecting relevant clusters, and then focusing on the specific laws within those clusters to determine the top five most appropriate principles. This approach resulted in a significant improvement in agreement between the SME and the chatbot. The taxonomy facilitated a more targeted and efficient search for applicable laws, leading to increased consistency in law selection and demonstrating its effectiveness in assisting both human experts and the chatbot in navigating the complex landscape of data science principles.

Significance of the Taxonomy

The specific categorization of the taxonomy may vary, but having a structured framework is crucial for matching laws to cases. The taxonomy organizes data science laws and principles logically, making it easier for users to identify relevant clusters and narrow down their search for applicable principles. This structured approach not only improves consistency in law selection between human experts and the chatbot, but also emphasizes the importance of having a clear and organized framework to guide decision-making in complex domains like data science management. The success of the taxonomy underscores the value of investing in the development of structured knowledge bases and decision-support tools to help professionals navigate vast amounts of information effectively.

Chatbot Validation Using Case Studies

To validate the effectiveness of the database of laws and the chatbot in a practical setting, five typical data science project management business use cases from business school were chosen (Agrawal, 2017), (Bizalwan, 2022), (Greenstein, 2018), (Lichtendahl, 2019), (Ovchinnikov, 2017). These cases were selected to exemplify common scenarios faced by industries where data science project management considerations are important. Each case presented unique challenges that tested the chatbot's ability to navigate the database of laws and provide accurate, context-specific recommendations. This step was crucial in demonstrating the chatbot's practical utility in real-world scenarios, echoing the project's goal to enhance the speed and accuracy with which professionals can access and apply security practices.

Integrating the database into the GPT model and the subsequent design of appropriate prompts were aimed at optimizing the chatbot's responses to be as tailored and relevant as possible. Initially, the model used only the compiled database of data science best practices. However, early trials revealed significant discrepancies between the GPT's responses and those of the SMEs, highlighting a gap in the model's ability to identify and apply the correct laws

autonomously. This issue was addressed by refining the taxonomy within the database, which enhanced the model's ability to generate more accurate and relevant recommendations.

To assess the chatbot's performance, a subject matter expert in data science was consulted. This expert, chosen for their extensive experience and current relevance in the field, provided authoritative answers and recommendations for each of the five selected business use cases. These expert insights served as a benchmark for evaluating the GPT model's outputs. The effectiveness of the chatbot and its underlying AI technology was quantitatively assessed using two statistical measures: the Confusion Matrix and Cohen's Kappa metric. The Confusion Matrix allowed for a detailed assessment of the model's correct and incorrect predictions, providing insight into the precision of the chatbot. Meanwhile, Cohen's Kappa metric offered a measure of the agreement between the chatbot's outputs and the expert's responses, accounting for the randomness that might influence such alignments.

This comprehensive evaluation approach was designed to test the feasibility of the proposed solution in generating relevant and compliant data science project recommendations. By leveraging an integrated knowledge base of laws and best practices, the research aimed to substantiate the chatbot's capacity to transform the landscape of data science management, directly addressing the challenges outlined in the introduction and supporting the broader goal of improving professional practices in the field of data science.

Chatbot Development

To create the DataSciLawBot, the GPT model was fine-tuned using a comprehensive dataset of curated data science laws and principles, structured into a taxonomy of nine thematic clusters. These clusters, which included areas such as model selection and optimization, data quality and preprocessing, and statistical principles, provided the framework for organizing the knowledge base. The fine-tuning process involved uploading this structured taxonomy along with the full set of data science laws into the GPT training environment. This ensured the model could access and understand a well-defined corpus of data science knowledge, enabling it to apply contextually appropriate laws to diverse scenarios.

The chatbot was programmed to initiate interaction by requesting specific details about the case or context in which the laws would be applied. For example, users could input descriptions of a data science project, highlighting its objectives, challenges, and constraints. Leveraging this input, the chatbot was designed to navigate the taxonomy, identify relevant clusters, and select the most applicable laws. It then provided detailed guidance on how those laws could be applied to address the problem or optimize the project. The responses included not only a justification for why particular laws were relevant but also practical recommendations tailored to the given scenario. This iterative process of fine-tuning and contextual problem-solving allowed the chatbot to serve as a dynamic and adaptive decision-support tool for data scientists, ensuring adherence to best practices in data science project management.

The chatbot is developed with the MyGPT platform, which is available in OpenAI's ChatGPT 4. It has been customized within ChatGPT to specifically cater to the needs of data science project management. This platform was chosen for its advanced natural language processing capabilities and its adaptability to specialized applications. It was fine-tuned with a detailed database of the laws presented in Appendix B. The prompt programming of the chatbot is documented in Appendix A. The GPT may be accessed here: <https://chat.openai.com/g/g-AKyo56VtB-datascilawbot>

Use Case

The following is a typical use case illustrating how this chatbot is meant to be used.

Emma, a senior data scientist, is leading a cross-functional team on a high-stakes project aimed at developing new machine learning models to predict market trends. The project involves multiple stakeholders, complex data integration from various sources, and tight deadlines.

User interface: Emma uses the DataSciLawBot to obtain insights on project management strategies specific to data science, such as handling data sprawl, ensuring data quality, and maintaining team productivity. **Specific Queries:** She asks about the best practices for managing large-scale data integrations and receives guidance on methodologies supported by data science laws, such as the principles of Data Lifecycle Management and Continuous Integration. **Outcome:** The advice from DataSciLawBot helps Emma streamline project workflows, improve team collaboration, and enhance the project's overall compliance with industry standards. The project culminates successfully with the models deployed ahead of schedule, significantly improving the company's analytical capabilities.

RESULTS

This study evaluated the feasibility of building an advanced chatbot powered by generative AI, specifically a GPT model, to generate relevant and accurate decision science recommendations based on a knowledge base of industry laws and best practices. The results demonstrate that while the initial implementation using only the raw database resulted in a significant gap between the chatbot's answers and those of subject matter experts (SMEs), implementing a taxonomy to group-related laws substantially improved the chatbot's ability to provide relevant and precise guidance. The chatbot's performance was assessed by comparing its responses to SME answers using Cohen's Kappa metric for inter-rater agreement and a Confusion Matrix for evaluating accuracy. The findings indicate substantial agreement between the chatbot and SMEs for individual law and cluster ratings, with more moderate alignment across the complete set of case-law pairs. This underscores the chatbot's potential as a decision support tool to augment human expertise while highlighting areas for further refinement.

Using the Chatbot

Figure 1 illustrates the chatbot's performance when analyzing a specific case study using the developed taxonomy and knowledge base. As shown in the screenshot, the chatbot was provided with a detailed description of a decision science scenario. Leveraging the structured database and its understanding of the relationships between various security principles, the chatbot processed the case details and generated a ranked list of the most relevant laws and best practices to consider. The output demonstrates the chatbot's ability to interpret complex, real-world situations and provide targeted, context-specific guidance. By identifying and prioritizing the key legal and industry standards applicable to the given case, the chatbot showcases its potential to support decision science professionals in navigating the complex landscape of cybersecurity regulations and best practices. The screenshot in Figure 1 offers a tangible example of how the chatbot can streamline the process of identifying and applying relevant security principles, ultimately enhancing decision-making and promoting more effective risk management strategies.

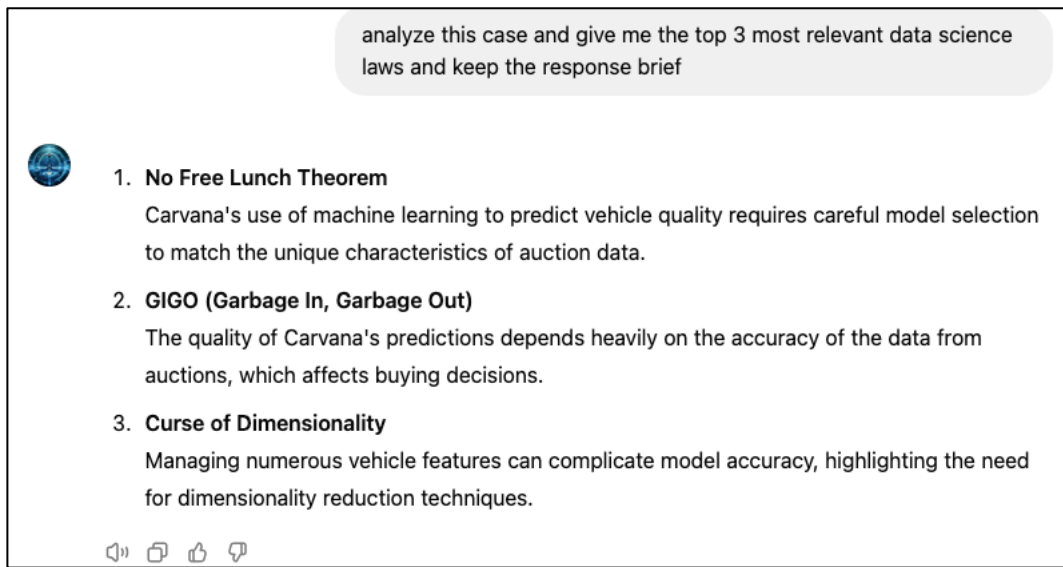


Figure 1: GPT prompt and results of analyzing a case using the chatbot

How to Use DataSciLawBot for Case Analysis

When faced with a case requiring guidance on the most relevant data science best practices, a data scientist can use the DataSciLawBot to ensure the project aligns with established industry standards.

The process begins with the data scientist providing a detailed description of the project or problem they are working on. This includes objectives, the type of data involved, the methodologies being considered, potential challenges, and any constraints or industry-specific concerns. The data scientist then enters the project details into the chatbot interface, posing specific queries such as "What best practices apply to managing unstructured data in this project?" or "How do I ensure alignment with data provenance principles?"

The chatbot uses the uploaded taxonomy of data science best practices to identify the clusters most relevant to the described context. For instance, a project dealing with cleaning and integrating datasets might trigger best practices within the "Data Quality and Preprocessing" cluster. Based on the identified clusters, the chatbot retrieves and ranks the most relevant best practices from the database. It provides tailored explanations of why each practice applies and how it can be implemented to address the specific challenges or goals outlined by the user.

If needed, the data scientist can refine their query or add more context to receive more targeted advice. For example, they might ask for alternative approaches or clarification on specific recommendations. The chatbot presents its findings, including practical steps to apply the identified best practices to the project. This ensures that the data scientist has a clear, actionable plan to follow, reducing the risk of oversight and enhancing project efficiency and adherence to industry standards.

By following this process, the DataSciLawBot empowers data scientists to manage their projects effectively while ensuring alignment with recognized best practices in the field.

Comparing SME and Chabot Analysis

We presented the case to both the SME and the chatbot. The SME was tasked with returning the top three laws as an ordered set. The chatbot was asked for the top five. The SME used the taxonomy as a decision tree to arrive at the answer. The selected laws by each contributor are shown in Figure 2 as ordered lists. It is interesting to note that in all cases but one, the top three laws selected by the SME are contained in the list of the top five laws selected by the chatbot. In the one case where the lists differ, they vary by just one law. We provide a more rigorous analysis below.

Case	Chatbot's Top 5 Laws	SME's Top 3 Laws
Case 1 BytelQ Analytics	Pareto Principle	If you're not paying, you're the product
	Principle of Least Effort	Pareto Principle
	Data Democratization	Principle of Least Effort
	Technical Debt	
	Algorithmic Fairness	
Case 2 Carvana	No Free Lunch Theorem	No Free Lunch Theorem
	Law of Large Numbers	Overfitting
	Data Provenance	Dijkstra
	Overfitting	
	Dijkstra's Principle	
Case 3 Scholastic Travel	No Free Lunch Theorem	No Free Lunch
	Law of Large Numbers	Lanchester's Laws
	Pareto Principle (80/20 Rule)	Human in the Loop
	Lanchester's Laws	
	Human in the Loop	
Case 4 Hospital Operations	Little's Law	Little's Law
	Correlation Precedes Causation	Correlation Precedes Causation
	Algorithmic Fairness	Algorithmic Fairness
	Data Provenance	
	Pareto Principle (80/20 Rule)	
Case 5 Viacom	Data Democratization	Data Democratization
	DEITT	DEITT
	Visualization and Tables	Visualization and Tables
	Feedback Loops	
	Continuous Integration	

Figure 2 –list of laws returned by the SME and the chatbot for all five cases.

Figure 3 illustrates the performance comparison between the DataSciLawBot chatbot and the subject matter expert (SME) assessment across five case studies. This matrix highlights the

agreement level between the chatbot and the SME in recognizing pertinent data science laws and principles for the top three laws identified by each expert, regardless of their ranking.

Case	Chatbot's Top 3 Laws	SME's Top 3 Laws	True Positives (TP)	False Positives (FP)	False Negatives (FN)
Case 1: BytelQ Analytics	Pareto Principle (80/20 Rule)	If you're not paying, you're the product	2	1	1
	Principle of Least Effort	Pareto Principle			
	Data Democratization	Principle of Least Effort			
Case 2: Carvana	No Free Lunch Theorem	No Free Lunch	1	2	2
	Law of Large Numbers	Overfitting			
	Data Provenance	Dijkstra			
Case 3: Scholastic Travel	No Free Lunch Theorem	No Free Lunch	2	1	1
	Law of Large Numbers	Lanchester			
	Human in the Loop	Human in the Loop			
Case 4: Hospital Operations	Little's Law	Little's Law	3	0	0
	Correlation Precedes Causation	Correlation Precedes Causation			
	Algorithmic Fairness	Algorithmic Fairness			
Case 5: Viacom	Data Democratization	Data Democratization	3	0	0
	DEITT	DEITT			
	Visualization and Tables	Visualization and Tables			

Figure 3 – Results of comparing SME and DataSciLawBot chatbot, which is the basis of a Cohen's Kappa test. Analysis of five case studies.

The results indicate that the chatbot performed as well as the SME in several key areas. Despite these differences, the overall agreement between the chatbot and the SME across clusters is noteworthy. Figure 3 underscores the chatbot's ability to provide recommendations that largely align with expert opinion. By considering a wide range of decision science laws and principles across multiple clusters, the chatbot demonstrates its capacity to offer comprehensive and nuanced guidance. The matrix highlights the potential for the chatbot to support and augment human expertise in analyzing complex cybersecurity scenarios and identifying relevant best practices.

While the chatbot's performance is promising, the discrepancies with the SME's assessment underscore the importance of further refinement and training to improve the chatbot's accuracy and consistency. Nonetheless, the results presented in Figure 3 provide compelling evidence for the chatbot's potential as a valuable tool in assisting data science professionals in navigating the complex landscape of data science laws and best practices.

	Predicted: True	Predicted: False
Actual: True	True Positives (TP) = 14	False Negatives (FN) = 4
Actual: False	False Positives (FP) = 4	True Negatives (TN) = 0

Figure 4 – Confusion matrix for the SME vs. Chatbot case study laws identification.

Figure 3 presents a confusion matrix that summarizes the overall performance of the DataSciLawBot chatbot compared to the subject matter expert (SME) in analyzing all the case

studies. The confusion matrix provides a comprehensive overview of the chatbot's ability to identify relevant decision science laws and principles, using the SME's assessment as the ground truth.

As shown in Figure 4, the chatbot achieved a significant number of true positives, indicating a high level of agreement with the SME in identifying relevant laws. The true negative count is also substantial, suggesting that the chatbot accurately recognized when certain laws were not applicable to the case. These results demonstrate the chatbot's ability to discern the relevance of decision science principles in the context of the case studies.

Outcomes

In the evaluation of the GPT model's alignment with the expert responses, the statistical measures provided insightful results, as illustrated in Figure 5, which summarizes the statistical tests for chatbot validation. The Cohen's Kappa metric, a statistical measure of inter-rater agreement, was applied across various aspects of the chatbot's output. For individual law ratings, Cohen's Kappa value reached 0.872, suggesting a substantial agreement between the GPT model's output and the subject matter expert's (SME) responses when the selections' order is not considered. We also see an F1 score of 93%.

Case	Chatbot's Top 3 Laws	SME's Top 3 Laws	True Positives (TP)	False Positives (FP)	False Negatives (FN)
Case 1: BytelQ Analytics	Pareto Principle (80/20 Rule)	If you're not paying, you're the product	2	1	1
	Principle of Least Effort	Pareto Principle			
	Data Democratization	Principle of Least Effort			
Case 2: Carvana	No Free Lunch Theorem	No Free Lunch	1	2	2
	Law of Large Numbers	Overfitting			
	Data Provenance	Dijkstra			
Case 3: Scholastic Travel	No Free Lunch Theorem	No Free Lunch	2	1	1
	Law of Large Numbers	Lanchester			
	Human in the Loop	Human in the Loop			
Case 4: Hospital Operations	Little's Law	Little's Law	3	0	0
	Correlation Precedes Causation	Correlation Precedes Causation			
	Algorithmic Fairness	Algorithmic Fairness			
Case 5: Viacom	Data Democratization	Data Democratization	3	0	0
	DEITT	DEITT			
	Visualization and Tables	Visualization and Tables			

Figure 5 – Details of the F1 score computation.

Metric	Value (nor order)	Value (pairwise)
True Positives (TP)	14	11
False Positives (FP)	4	4
False Negatives (FN)	4	4
Precision	0.933 (93.3%)	0.733 (73.3%)
Recall	0.933 (93.3%)	0.733 (73.3%)
F1 Score	0.933 (93.3%)	0.733 (73.3%)
Cohen's Kappa (κ)	0.872	0.567

Figure 6 – Summary of statistical tests for chatbot validation.

The laws in Figure 4 are shown in the order that the chatbot and the SMEs returned from most applicable to least. The results provided thus far are for computations without concern for order. Going further and analyzing the results for pairwise comparison of the ordered returns, we obtain slightly lower values, a Cohen's Kappa of .567 and an F1 score of 73%, tabulated in Figure 6.

We want to emphasize that while this is a useful measure, it is not perfect. This analysis relies on a small dataset of just five cases, which limits the validity of the statistical outcomes—such as Precision, Recall, F1 Score, and Cohen's Kappa. The restricted sample size makes the results particularly sensitive to minor changes, undermining their statistical reliability and applicability. As a result, the metrics observed may not accurately reflect the chatbot's overall performance due to the narrow focus, indicating they are more suggestive than definitive. To draw stronger and more reliable conclusions, it is essential to utilize a larger and more varied dataset.

Implications

The research aims to explore how generative AI models can be used to help with complex decision-making processes, especially in decision science. The proposed approach could be used as a decision support tool to offer initial recommendations based on relevant laws and best practices. These recommendations can then be reviewed and improved by human experts.

Evaluation of Research Questions

Research Question 1 – How feasible is it to build a generative AI-powered chatbot that can accurately deliver context-specific guidance on decision science best practices to professionals in real time?

The feasibility of constructing a generative AI-powered chatbot that provides context-specific guidance on decision science best practices has been largely demonstrated by the substantial agreement observed in Cohen's Kappa values for both law and cluster ratings. The values of 0.872 indicates that the chatbot can provide responses that align significantly with the insights of a subject matter expert (SME). Additionally, the accuracy reflected by an F1 score of 0.933 for law ratings further supports the chatbot's ability to deliver relevant and precise guidance. These metrics show that while there are areas for improvement, the chatbot effectively interprets and applies information from a comprehensive database of security best practices in a manner that is both timely and contextually appropriate. Thus, building and implementing such a chatbot for real-time professional use is feasible.

Research Question 2 – How well will such a chatbot perform compared to human SMEs?

A comparison of the chatbot's performance with that of human subject matter experts (SMEs) presents a mixed yet encouraging outcome. For specific laws and groups of information, the chatbot shows a notable alignment with SMEs, as demonstrated by the high values of Cohen's Kappa. This indicates that in certain contexts where precise and targeted guidance is necessary, the chatbot can compete with human knowledge. However, when assessing the performance over the complete range of case-law pairs, Cohen's Kappa value falls to 0.567, reflecting only moderate agreement. This difference suggests that while the chatbot excels in controlled or particular situations, its capacity to consistently emulate the nuanced insight of human experts across a wider array of scenarios is less dependable.

This variance indicates AI's current limitations in handling complex, multifaceted queries that may require deeper insight or a more holistic understanding than what is currently achievable through automated means. Nonetheless, the chatbot does present a valuable tool for augmenting human capabilities, providing quick, initial assessments that can then be further refined or validated by human experts.

Hypothesis Testing

H1: An LLM-based chatbot fine-tuned with decision science laws will apply these laws to solve data science cases just as well as a human subject matter expert.

The substantial agreement indicated by Cohen's Kappa value of 0.872 for unpaired comparison) supports the hypothesis that the chatbot significantly enhances the accuracy of accessing and applying best practices. The high F1 score (0.93) further validates the chatbot's effectiveness, though it also highlights areas for improvement in decision-making accuracy. The fair agreement for the entire set of case-law pairs (Cohen's Kappa value of 0.567) indicates that while the chatbot performs well on individual components, achieving comprehensive agreement across complex case-law scenarios remains challenging.

Summary of Results

This study evaluated the feasibility of building an advanced chatbot using a generative pre-trained transformer (GPT) model to generate relevant and compliant decision science recommendations based on a knowledge base of laws and best practices. Initially, the GPT model used only the collected database of 53 decision science best practices. However, the results showed a significant gap between the GPT's answers and subject matter experts' (SMEs) answers, as the GPT model attempted to find laws by itself. To address this, a taxonomy was implemented to group the laws, improving the GPT model's ability to generate more accurate and relevant recommendations.

CONCLUSIONS AND FURTHER WORK

The feasibility study, which involved building and testing the DataSciLawBot GPT, demonstrated its potential to facilitate access to critical information and provide decision-making support for data science professionals. This outcome emphasizes the practical benefits of AI in enhancing professional efficacy. The effectiveness of the GPT bot validates the concept that AI can significantly improve the efficiency of accessing and utilizing industry best practices. However, the success of such tools is heavily dependent on the precision of the input data and the clarity of user interactions.

The proof of concept demonstrated that when properly configured, the GPT could effectively assist in complex decision-making environments. This was demonstrated by its ability to quickly provide relevant and accurate information. This project adds to the ongoing discussion about AI applications in specialized fields, offering a case study on customizing AI tools to meet specific professional needs. For data science practitioners, the GPT bot represents a significant improvement in navigating and implementing best practices, potentially reducing the time and effort needed for manual research and analysis.

The project highlighted the difficulty of programming AI to consistently process highly technical content. It emphasized the importance of sophisticated natural language understanding capabilities. Concerns about AI biases were addressed by emphasizing the need for diverse and comprehensive datasets to train the model, ensuring that it provides balanced and impartial advice.

In the future, we plan to improve the AI's algorithms to better handle unclear questions and expand the training dataset. In the short term, we will conduct an experimental phase to compare answers from information security students and industry employees with those generated by GPT. This will provide insights into whether the tool saves time and delivers accurate advice.

In the long term, we could explore using this GPT bot in other sectors within the information security industry or adapt the model for other fields requiring specialized knowledge management. This project has shown that with further improvements, advanced AI tools like GPT can support information security professionals by providing accurate guidance and streamlining access to industry best practices. Future evaluations will focus on the bot's effectiveness and efficiency to ensure practical utility in real-world scenarios.

REFERENCES

- Agrawal, S. (2017). *Why hospitals need better data science*. Harvard Business Review.
- Battineni, G., Chintalapudi, N., & Amenta, F. (2020). Ai chatbot design during an epidemic like the novel coronavirus. *Healthcare (Basel)*, 8(2), 154. <https://doi.org/10.3390/healthcare8020154>
- Bizalwan, D. S., Chauhan, S., & Bajaj, G. (2022). *BytelQ: Service vs. product or survival vs. democratization?* Ivey Publishing.
- Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success. *Information Systems Frontiers*, 25(1), 161–182. <https://doi.org/10.1007/s10796-022-10291-4>
- Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., & Tian, F. (2021). Utilization of self-diagnosis health chatbots in real-world settings: Case study. *Journal of Medical Internet Research*, 23(1), e19928. <https://doi.org/10.2196/19928>
- Ferreira, B., Marques, S., Kalinowski, M., Lopes, H., & Barbosa, S. D. J. (2023). Lessons learned to improve the UX practices in agile projects involving data science and process automation. *Information and Software Technology*, 155, 107106. <https://doi.org/10.1016/j.infsof.2022.107106>
- Fortino, A. and Zhang, W. (2024), AI-Enhanced Decision-Making: Integrating It Best Practices With Expert Insight, Decision Sciences Institute Conference, Phoenix, AZ Nov. 21-23, 2024.
- Greenstein, S., & Snively, C. (2018). *Viacom: Democratization of data science*. Harvard Business School.
- Hassani, H., & Silva, E. S. (2023). The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data and Cognitive Computing*, 7(2), 62. <https://doi.org/10.3390/bdcc7020062>
- Kintsakis, A. M., Psomopoulos, F. E., & Mitkas, P. A. (2019). Reinforcement Learning based scheduling in a workflow management system. *Engineering Applications of Artificial Intelligence*, 81, 94–106. <https://doi.org/10.1016/j.engappai.2019.02.013>
- Lichtendahl, K. C. Jr., & Holtz, P. (2019). *Carvana: IsBadBuy?* University of Virginia Darden School Foundation.
- Luo, B., Lau, R. Y. K., Li, C., & Si, Y. (2022). A critical review of state-of-the-art chatbot designs and applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12(1), e1434. <https://doi.org/10.1002/widm.1434>
- Ovchinnikov, A. (2017). *Retention modeling at Scholastic Travel Company (A)*. University of Virginia Darden School Foundation.
- Santos, G. A., de Andrade, G. G., Silva, G. R. S., Duarte, F. C. M., Costa, J. P. J. D., & de Sousa, R. T. (2022). A Conversation-Driven Approach for Chatbot Management. *IEEE Access*, 10, 8474–8486. <https://doi.org/10.1109/ACCESS.2022.3143323>
- U.S. Bureau of Labor Statistics. (2019). Employment projections—2019-2029.
- U.S. Bureau of Labor Statistics. (n.d.). Occupational Outlook Handbook: Data Scientists. Retrieved from <https://www.bls.gov/ooh/>
- Wilkinson, G. G. (2024). Enhancing Generic Skills Development in Higher Education in the Era of Large Language Model Artificial Intelligence. *Journal of Higher Education Theory and Practice*, 24(3), 64–76. <https://doi.org/10.1080/10691898.2020.1860725>
- Yu, Y., Fortino, A. and Felberbaum, A. (2024), InfoSecPilot: Navigating the Complex Landscape of Information Security with an AI-Powered Knowledge Management Chatbot, Decision Sciences Institute Conference, Phoenix, AZ Nov. 21-23, 2024.

APPENDIX A

Chatbot Prompts

To Develop a Taxonomy

PROMPT

You are a chief AI officer with many years of experience in computer science and in data science. You're interested in compiling a set of industry best practices and laws that can guide managing data, science, projects and investments in data science activities for your company. You want to use this for yourself and for your managers. You assign one of your staff to compile these laws, and they come back with these laws and principles found in the uploaded document. Now with your wisdom and experience, you're going to group these laws in clusters, no more than 12, but no less than seven at your discretion. Each law is placed into one of these clusters. Then, provide a list of the clusters with the name, a short description of the cluster, as well as which laws are assigned to each cluster in your opinion.

APPENDIX B

Data Science Laws and Principles

1. Anscombe's Quartet: "Identical summary statistics can come from vastly different data." Look at the raw data, not just the summary results.
2. Bayes' Theorem: "Describes the probability of an event, based on prior knowledge of conditions that might be related to the event." Useful for inferential statistics and machine learning.
3. Bayes' Theorem: Provides a mathematical framework for updating probability estimates based on new evidence and prior understanding. Essential foundation for statistics and machine learning approaches.
4. Benford's Law: The observation that in many real-world data sets, the leading digit is likely to be small. This can be used to check for fraudulent or randomized data. Useful as a basic sanity check before further analysis.
5. Central Limit Theorem: The mean of sampled subsets approaches a normal distribution irrespective of the distribution of the overall population. Enables use of statistical testing based on normality assumptions.
6. Concept Drift: "The relationship between variables can change over time." Analytics systems need robustness to adaptation.
7. Confirmation Bias: "Tendency to favor information confirming preexisting views." Deliberately challenge assumptions.
8. Continuous Integration: "Frequently merging updated code enables identifying conflicts sooner." Avoid big bang risky releases.
9. Correlation Precedes Causation: "A relationship between two variables does not necessarily imply that one causes the other." Experimentation is required to confirm causal links.
10. Dark Data: "Underutilized information wastes value." Discover and apply overlooked data sources.
11. Data Democratization: "Information should be accessible to all stakeholders to empower decision making", But balance openness with governance.
12. Data Drift: "Populations change over time so models need periodic refreshing." Ensure monitoring and adaptation.
13. Data Ethics: "Data collection and usage should respect human rights and privacy." Ensure transparency, consent and avoiding bias.
14. Data Provenance: "Understanding the origins and processing history of data is essential for evaluating and using it properly." Metadata provides context for appropriate interpretation.
15. Data Rich, Information Poor: "Collecting increasing data outpaces gaining useful knowledge from it." Focus on information not just data.
16. Data Velocity: "Rate of data generation is accelerating exponentially." Architectures must handle speed and volume.
17. Datasets Ending in Tears (DEITT): "Just because data is used in analysis does not automatically make decisions better." Numbers supplement but cannot replace judgement.
18. Fail Fast: "Learning happens faster when we quickly try imperfect solutions, fail, and improve." But balance speed with thoroughness when needed.
19. Feedback Loops: "Outcome changes system, which then affects future outcome." Understand cascading and compounding effects.

20. Fourier Transform: Allows transformation of signals between time and frequency representations. Enables spectral analysis and filtering for time series data. Underpins signal processing techniques.
21. Fundamental Attribution Error: "Overemphasizing personalities versus contextual factors when judging behavior." Consider external influences on observed outcomes.
22. Garbage In, Garbage Out: "Flawed input data produces worthless output." Ensuring data quality and cleansing is an essential first step for analysis.
23. Gradient Descent Optimization: Iterative method to minimize differentiable loss functions by taking partial derivative steps proportional to slope. Core principle powering neural network learning processes.
24. Heaps' Law: The number of distinct words in a document grows according to a sub-linear power law function of the document length. This relationship allows estimating the number of new topics/concepts likely to emerge in longer texts.
25. If You're Not Paying, You're the Product: "When services are free, user data is the real business model." Understand the behind-the-scenes quid pro quo.
26. Information Entropy: Quantifies the unpredictability or information content of messages or datasets. Measures disorder and uncertainty to highlight areas needing more clarification.
27. Law of Large Numbers: "Increasing the sample size improves estimate accuracy and consistency." But more data isn't always better or possible.
28. Law of Large Numbers: As sample size grows, its mean gets closer to the population mean. Justifies reliance on statistics derived from big datasets.
29. Linked Data: "Connecting related data creates more value than silos." Combine complimentary datasets when possible.
30. Loose Tight Coupling: "Components should have little dependency on each other but still work together well." Enables isolating failures and independent updates.
31. Map-Reduce: "Large datasets handled efficiently by distributing work across clusters of machines." Enables massively parallel processing.
32. Models vs Explanations: "Predictive models describe what not necessarily why." Explanation requires additional investigation.
33. No Free Lunch Theorem: "There is no one superior machine learning algorithm for all cases." Understand tradeoffs to match approach to problem.
34. Occam's Razor: "Simpler solutions are more likely to be correct than complex ones." Preference for simplicity and elegance should guide model selection.
35. Overfitting: "Models that are excessively complex relative to the amount of data may capture noise rather than the underlying relationship." Avoid overfitting with regularization, cross-validation etc.
36. Pareto Principle: 80% of effects come from 20% of the causes. This highlights the vital few x-factors versus the trivial many. Helpful for identifying the key drivers and phenomena to focus analysis on first.
37. Price's Law: The square root of the number of people in a domain do 50% of the work. Highlights the asymmetric nature of contribution even among subject experts. Also indicates returns possible by engaging more contributors.
38. Principal Component Analysis (PCA): Method to reduce high dimensionality data into fewer dimensions capturing maximal variance. Useful for feature extraction, noise reduction, and visualization.
39. Propagation of Uncertainty: Mathematical techniques for rigorously tracking how uncertainties in input variables translate into confidence bounds around calculated outputs. Critical for trustworthy ML predictions.
40. Proxy Variables: "Indirect measurements that correlate to unmeasurable items can provide useful approximations of them." But they may include unknown biases.

41. Regression to the Mean: "Unusually large or small numbers are likely to move towards average over time." Do not mistake outliers or random fluctuations for trends.
42. Sampling Bias: "Non-random data collection may not represent the larger truth." Beware systemic distortions and find more representative samples.
43. Signals and Noise: "Relevant data tends to be drowned out by irrelevant variability." Statistical techniques needed to reveal insights.
44. Simpson's Paradox: "Trends can reverse when different groups are combined." Deeper analysis is needed to understand the true relationships.
45. Small Data: "Simple cases reveal powerful insights." Start simple before tackling entire systems.
46. Stigler's Law of Eponymy: Scientific discoveries are rarely named after their actual discoverer. Useful reminder that it's the knowledge that matters more than the credit. Data science should focus on collaboratively expanding collective understanding.
47. Streaming over Batch Processing: "Real-time analysis enables instant learning and intervention." But batch processing easier for complex or historical analysis.
48. Technical Debt: "Expedient solutions incur additional rework cost later, like debt financing." Refactor and optimize code for sustainability.
49. The Multiple Comparisons Problem: "Performing many tests inflates the chances of false positives." Apply techniques like Bonferroni adjustment or Benjamini-Hochberg procedure for multiple testing.
50. The Pareto Principle: "80% of outcomes come from 20% of causes." Identify and focus on the vital few drivers rather than trivial many.
51. Visualization over Tables: "Charts and graphs facilitate understanding complex data far better than raw numbers alone." Exploit human's visual perception for insight.
52. Wisdom of Crowds: "Large groups can predict more accurately than individuals." But crowds are fallible and other viewpoints should still be considered.
53. Zipf's Law: The n th most common item occurs about $1/n$ times as often as the most common item. Useful for predictive analytics and modeling language and information distributions. Provides insight into the long tail as well.
54. JT Datafication: "Quantifying phenomena facilitate data-driven decisions about them, but risks oversimplification." Balance quantification with broader context.

APPENDIX C

Taxonomy of Data Science Laws

Cluster	Description	Laws Assigned
Model Selection and Optimization	Best practices for choosing and fine-tuning models effectively.	Occam's Razor, Overfitting, No Free Lunch Theorem, Gradient Descent Optimization, Models vs Explanations, Concept Drift
Data Quality and Preprocessing	Ensuring data quality and readiness for analysis.	Garbage In, Garbage Out, Data Provenance, Sampling Bias, Data Ethics, Data Drift, Data Rich Information Poor
Statistical Principles and Interpretation	Foundational statistics for data-driven insights.	Bayes' Theorem, Law of Large Numbers, Central Limit Theorem, Regression to the Mean, The Multiple Comparisons Problem, Benford's Law, Zipf's Law, Heaps' Law
Causal Analysis and Relationships	Understanding the true relationships between data points.	Correlation Precedes Causation, Simpson's Paradox, Proxy Variables, Feedback Loops
Human Judgment and Cognitive Biases	Awareness of human factors in interpreting data.	Confirmation Bias, Fundamental Attribution Error, Datasets Ending in Tears (DEITT), Data Ethics (cross-referenced), If You're Not Paying, You're the Product
Data Analysis and Visualization	Emphasizing effective data exploration and communication.	Anscombe's Quartet, Visualization over Tables, Signals and Noise, Pareto Principle, Wisdom of Crowds
Data Engineering and Architecture	Infrastructure for efficient data processing and analysis.	Map-Reduce, Streaming over Batch Processing, Loose Tight Coupling, Data Velocity, Linked Data
Agile Development and Technical Debt	Managing development processes and avoiding costly	Fail Fast, Technical Debt, Continuous Integration, Dark Data
Uncertainty and Complexity Management	Handling the unpredictable aspects of data science.	Information Entropy, Propagation of Uncertainty, Small Data, Price's Law, Stigler's Law of Eponymy

Design Science and Behavioral Science in Information Systems Research
(From 2004 to 2024)

Toyin Asonibare

Department of Information Science & Systems

Morgan State University

Baltimore, MD 21251

Email: toasol@morgan.edu

Ganesh Bhatt

Department of Information Science & Systems

Morgan State University

Baltimore, MD 21251

Email: ganesh.bhatt@morgan.edu

Abstract

Design science and behavioral science are two research paradigms that underpin most of the Information Systems (IS) research. They both focus on the core element of IS i.e. technology and the behavior of people when interacting with technology. Design science focuses on developing artifacts to solve organizational problems. Behavioral science focuses on the development and justification of theories that account for human and organizational behaviors. Design science research has gained popularity among IS scholars since the publication of *Design Science in Information Systems Research* by Hevner, March, Park, and Ram in 2004. Some scholars have emphasized the complementary use of both paradigms for rigor and meaningful contribution to IS research. This study explores how IS scholars engage with design science research and examines their level of engagement with the complementary use of design science and behavioral science paradigms. A bibliometric analysis was performed on papers published in *Decision Support Systems (DSS)*, *European Journal of Information Systems (EJIS)*, *Information Systems Journal (ISJ)*, *Information Systems Research (ISR)*, *Journal of the Association for Information Systems (JAIS)*, *Journal of Information Technology (JIT)*, *Journal of Strategic Information Systems (JSIS)*, *Journal of Management Information Systems (JMIS)*, *MIS Quarterly (MISQ)*, *Information and Management (IM)*, the *Americas Conference on Information Systems (AMCIS)*, the *International Conference on Information Systems (ICIS)*, and the *International Conference on Design Science Research in Information Systems and Technology (DESRIST)*, from 2004 to 2024.

Our findings reveal that design science research has matured and is now widely recognized as a key paradigm in the IS field, with most publications originating from conference outlets. However, while the complementary use of design science and behavioral science paradigms is

acknowledged in the literature, the level of engagement of IS scholars with the complementary use of these paradigms remains notably low.

Keywords: Design Science, Design Research, Design Science Research, Behavioral Science, Information Systems, Bibliometric Analysis

1. Introduction

The field of Information Systems (IS) is characterized by diversity in research problems, methods, and theoretical foundations (Ayanso et al., 2007). This diversity arises from the various reference disciplines that make up the IS field. Scholars working in IS bring along with them the methodology and theories in which they are trained (Vessey et al., 2002). Given the adoption of different approaches in IS research, some scholars have emphasized the importance of rigor (Benbasat & Weber, 1996; Hevner et al., 2004), while others have advocated for the need for multiple theories and the adoption of the appropriate paradigms to address diverse research issues in the field (e.g. Banville & Landry, 1989; Huysmans & De Bruyn, 2013).

Design science and behavioral science are two such paradigms that underpin much of IS research (Deng & Ji, 2018; Hevner et al., 2004). Both paradigms focus on the core elements of the field i.e. technology and the behavior of people when interacting with technology (Lukyanenko, 2021), and play a pivotal role in the development of collective knowledge that shapes the IS field (Niederman & March, 2012).

Design science originates from engineering (March & Smith, 1995). It is concerned with how IT artifacts are developed to solve organizational problems (Huysmans & De Bruyn, 2013).

Behavioral science, on the other hand, originates from natural science (Hevner et al., 2004). It is

concerned with the development and justification of theories that account for human and organizational behavior (Ayanso et al., 2011; Hevner et al., 2004). The goal of design science is to develop an IT artifact that is useful to address real-world problems, while behavioral science seeks truth that explains use behavior (Goldkuhl, 2016; March & Smith, 1995).

Although design science and behavioral science are two different paradigms, they complement each other in IS research (Lukyanenko, 2021; Wimmer & Yoon, 2017). Their complementary use allows IS scholars to address research problems comprehensively by not only focusing on the creation of useful artifacts to solve organizational problems but also evaluating their effect on users and organizations (Sutton et al., 2021). This approach enhances the efficiency and relevance of IS research (Hevner et al., 2004) and enables scholars to make meaningful contributions to the field (Niederman & March, 2012).

This study aims to achieve two key objectives: first, to examine the current state of design science research in the IS field and how IS scholars are engaging with design science in their research, and second, to explore the extent of the complementary use of design science and behavioral science paradigms in IS research. To address these objectives, the following research questions are posed:

- **RQ1:** To what extent are IS scholars engaging with research pertaining to design science?
- **RQ2:** To what extent are IS scholars engaging with the complementary use of design science and behavioral science in their research?

This study contributes to the existing body of knowledge by applying bibliometric analysis to intellectually map design science research in the IS field. Specifically, it adds to the discourse on

the engagement of IS scholars with design science research and the complementary use of design science and behavioral science paradigms within the field. By analyzing design science research publications over a period of twenty years, from 2004 to 2024, this study provides insights into the level of engagement of IS scholars with these paradigms in their research.

The findings of this research will provide scholars with a clearer understanding of the extent to which design science and the integration of design science and behavioral science are represented in IS research. This is important because the integration of these paradigms offers a more holistic approach to addressing IS problems (Hevner et al., 2004) and fosters the rigorous and relevant solutions sought by the IS community (Niederman & March, 2012). The remaining part of this paper is organized as follows: Section 2 provides a literature review on design science and behavioral science research in IS. Section 3 outlines the research methodology employed in this study. Section 4 presents the findings and discussion, and Section 5 concludes the paper with suggestions for future research directions.

2. Literature Review

2.1. Design Science Research in IS

Behavioral science has traditionally been the dominant paradigm among IS scholars (Huysmans & De Bruyn, 2013; Jain et al., 2011). This research stream focuses on phenomena that emerge when individuals interact with IT artifacts introduced within an organization (Hevner et al., 2004). Design science research, on the other hand, has gained increasing prominence in the IS community in recent years (Goldkuhl, 2016), particularly following the seminal publication of *Design Science in Information Systems Research* by Hevner et al. (Metallo et al., 2012). Their work highlighted the importance of design science and provided IS scholars with guidelines for

conducting design science research. The growing recognition of design science research in the IS community is evident in developments such as special issues dedicated to design science in leading IS journals like MIS Quarterly (Goes, 2014), and the establishment of the Conference on Design Science Research in Information Systems and Technology (DESRIST) in 2006 (Fischer, 2011). These initiatives offer avenues for the growth of design science research within the IS community.

Baskerville (2008) contends that there is a lack of consensus on what design science is. He addressed this by delineating what design science is not. The author examined the selected best papers at the 2008 DESRIST conference. Based on the insights drawn from these papers, Baskerville argues that design science should not be confused with related constructs such as design, design theory, action design, IT artifacts, computer science, or methodology, nor should it be regarded as a separate academic discipline. Thuan et al. (2019) recognize the lack of guidance on how to effectively frame research questions in design science research. The authors analyzed research questions from 104 publications on design science from DESRIST, HICSS, MISQ, and PhD dissertations. They observed that most publications constructed their research questions by focusing on identified problems, while others used gap spotting and problematization. The authors proposed a typology to guide scholars in formulating research questions for their design science research. Antunes et al. (2021) conducted a systematic literature review of 54 design science conceptual frameworks published in the Association for Information Systems Electronic Library (AISEL) to explore the use of conceptual frameworks in design science research. They found that conceptual frameworks in design science facilitate “problematizing, positioning the research, organizing the inquiry process, representing DSR artifacts, and explaining design processes” (p. 82). The authors developed a decision model to guide researchers in creating

conceptual frameworks. This model takes into account the research stage, the researcher's perspective, and the dimensions of the study. Maedche et al. (2021) argue that IS scholars lack an adequate understanding of what design science contribution entails and the various research activities that give rise to this contribution. The authors proposed a Design Research Activity framework. This framework categorizes design science contributions into the following segments: "construction, manipulation, deployment, and elucidation" (p.14). The authors applied this framework to categorize design science papers in the Journal of the Association for Information Systems (JAIS) from 2007 to 2019 across the four segments. They found that design science research is distributed across the four segments and that design science research comprises design science and design-oriented behavioral research. Deng et al. (2017) contend that design science research has gained recognition as a complementary research paradigm to behavioral science and is widely used to address various IS topics. Through a systematic literature review, the authors examine the current state of design science research in the IS field. Analyzing 278 design science research articles published between 2001 and 2015 in top IS journals and conference proceedings, they found that most publications focused on the methodologies for conducting design science research and the evaluation of artifacts. Baskerville et al. (2018) argue that, while design science has gained popularity in the IS field, its publication rate in top journals remains relatively low. The authors attribute this to the theoretical contribution required of design science research by IS journal reviewers. To address this issue, they provided guidelines for researchers and reviewers outlining what is required for design science research to be accepted for publication to increase the acceptance rate of design science research in top IS journals.

The literature reflects the significant contributions of IS scholars to the growth and acceptance of design science research within the academic community. Some scholars, such as Baskerville (2008), focus on conceptualizing design science research, while others, including Baskerville et al. (2018), address challenges like the low acceptance rates of design science research in top IS journals. Scholars such as Maedche et al. (2021), Thuan et al. (2019), and Antunes et al. (2021) provide guidelines for categorizing design research contributions, framing design science research questions, and creating conceptual frameworks in design science research. Also, the work of Baskerville et al. and Maedche et al. serves as a guide for both researchers and reviewers in identifying design science research contributions, thereby enhancing its acceptance rate.

2.2. Design Science and Behavioral Science Research in IS

From its inception, design science has been acknowledged as complementary to behavioral science (Huysmans & De Bruyn, 2013). Hevner et al. (2004) emphasizes the inseparable relationship between technology and behavior in IS research. They call for a complementary effort between design science and behavioral science scholars.

Niederman and March (2012) recognize the growing importance of design science in IS research. The authors argue that addressing sociotechnical issues in IS requires the integration of behavioral science and design science paradigms. They examine the relationship between design science and behavioral science and delineate the importance of combining these paradigms in research. Niederman and March are of the view that behavioral research should be integrated in the design and evaluation of IT artifacts for effectiveness and relevance. Lukyanenko (2021) is of the view that design science and behavioral science research are two distinct but complementary paradigms in the IS field. Despite their potential for synergy, these communities rarely collaborate. To address this gap, Lukyanenko proposes a novel approach termed Design

Behavioral Research (DBR) and introduces Type VI theory, described as the “theory of explanation, prediction, and action”(p.1), which provides a theoretical foundation for integrating these paradigms in IS research. Similarly, Huysmans and De Bruyn (2013) contend that the complementary use of design science and behavioral science paradigms in IS research is acknowledged in the literature, but there are no guidelines on how to integrate them. To address this, the authors propose a mixed-method approach as a guide for integration. They identified five types of mixed methods: concurrent exploratory, concurrent creative, concurrent evaluative, sequential confirmatory, and sequential explanatory. These methods enable the integration of behavioral research within design research. Goldkuhl (2016) argues that design science and behavioral science are often seen as distinct paradigms in the IS field. The author proposed a unifying perspective of the two paradigms which he refers to as post-hoc studies (behavioral research) and design studies of IS research. Goldkuhl is of the view that the unified perspective fosters cross-fertilization among researchers and promotes a more comprehensive understanding of IS design and behavioral research. Wimmer and Yoon (2017) contend that IS scholars have traditionally treated design science and behavioral science as separate paradigms. The authors advocate for the complementary use of these paradigms. They illustrate the integration of design science and behavioral science through the development of the OnCDS system, an artifact designed to detect counterfeit products by analyzing online reviews. They evaluate the artifact’s impact on purchaser attitudes using a model incorporating trust and perceived risk. Their study demonstrates how the integration of design science and behavioral science can enhance artifact design and the understanding of user behavior. Sutton et al. (2021) argue that design science research has traditionally focused on artifact creation, often overlooking the role of behavioral science in evaluating how these artifacts are used by humans.

To address this, the authors extend the design science framework to include behavioral evaluation. They used the INSOLVE Research Program to illustrate the complementary relationship between design science and behavioral science, where each paradigm informs and enhances the other.

The reviewed literature underscores the recognition of the complementary use of design science and behavioral science paradigms in IS research. While Goldkuhl (2016) challenges the separation of these paradigms by IS scholars, Lukyanenko (2021), Huysmans and De Bruyn (2013), and Niederman and March (2012) propose frameworks for their integration to address issues in the IS field. The development and evaluation of the OnCDS by Wimmer and Yoon (2017) demonstrates how these paradigms can be effectively combined in practice. Similarly, the INSOLVE program presented in Sutton et al. (2021) shows the iterative synergy between artifact creation in design science and behavioral science evaluation, demonstrating how these paradigms can inform each other.

There have been calls for the integration of design science and behavioral science research in the field of IS (e.g. Goldkuhl, 2016; Hevner et al., 2004; Kelton & Murthy, 2023) and efforts have been made by some scholars to explore the complementary use of these paradigms (e.g. Sutton et al., 2021; Wimmer & Yoon, 2017). In addition, some guidelines have been proposed to facilitate the integration of these paradigms (e.g. Huysmans & De Bruyn, 2013; Lukyanenko, 2021; Sutton et al., 2021). However, no existing work to the best of our knowledge has examined the extent to which IS scholars, in particular design science researchers, are engaging with the complementary use of these paradigms in their research. This represents a critical gap in the literature and underscores the need for further research to evaluate the complementary use of these paradigms in IS research.

3. Research Methodology

3.1. Overview of Bibliometric Analysis

Our research methodology is based on bibliometric analysis. Bibliometric analysis is a scientific research method that is used for quantitative analysis of literature (Donthu, Kumar, Mukherjee, et al., 2021; Pascal & Renaud, 2020). Bibliometric analysis has become increasingly popular as a scientific tool, widely utilized by scholars for two main purposes; performance analysis and science mapping (Donthu, Kumar, Mukherjee, et al., 2021). It has been employed to explore the intellectual structure of a field (e.g. Culnan, 1986), analyze the patterns of scholarly communication within a community (e.g. Pascal & Renaud, 2020), and identify potential directions for the future development of a field (e.g. Walsh & Renaud, 2017). Bibliometric analysis can be conducted using five distinct methods: citation analysis, co-citation analysis, bibliographic coupling, co-authorship analysis, and co-word analysis (Zupic & Čater, 2015). Metallo et al. (2012) perform a bibliometric analysis of 245 publications spanning 1995 to 2012 to examine the state of design science research within the IS field. Their study analyzed publications from seven major IS conference outlets, including ACIS, AMCIS, ECIS, ICIS, PACIS, HICSS, and DESRIST, as well as IS journals. The findings revealed a steady increase in design science publications across conferences and journals, particularly after the publication of Hevner et al. (2004) seminal work, which has significantly influenced the field. Their contributions addressed various aspects such as guidelines for conducting design science research, and its practical applications. Pascal and Renaud (2020) examine the evolution of IS design science research over a 15-year period following the seminal work of Hevner et al. (2004). By analyzing 192 design science research articles from the AIS basket of eight journals,

the authors used co-citation analysis to identify five intellectual clusters central to design science research: qualitative methodological foundations, reflections on the IS discipline, IS design theory, action research in IS design science, and general guidelines for IS design science research. Based on these clusters, they proposed an integrative framework to help IS researchers position their work within the broader intellectual landscape. Dwivedi et al. (2010) employ citation analysis to compare the engagement of IS scholars with the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT). Their findings reveal similarities between the two models, with UTAUT gaining increasing attention among researchers. The authors emphasize that this shift necessitates a rigorous evaluation of studies employing UTAUT to ensure originality and to safeguard against a narrowing of knowledge diversification within the field. Fischer (2011) examine the evolution of design science research in MIS by analyzing 50 design science articles published between 2007 and 2010 in six leading IS journals, MISQ, ISR, JMIS, JAIS, EJIS, and ISJ. The findings of the citation analysis indicate that 70% of these articles referenced Hevner et al. (2004) seminal work, demonstrating the cumulative tradition of design science research and its foundational role in the field. In celebration of the 40th anniversary of the International Journal of Information Management (IJIM), Donthu, Kumar, Pandey, and Gupta (2021) reflect on the journal's contributions through bibliometric analysis. Their study highlights an increase in publications and citations over the years, with a majority originating from Europe and the United States. They found that most articles are empirical studies employing quantitative methodologies. Their research contributes to the understanding of IS journal evolution, offering insights into key trends, theoretical advancements, and patterns of scholarly communication within the field.

The bibliometric analysis for this study is conducted using VOSviewer version 1.6.20.

VOSviewer is freely available software designed for creating and visualizing bibliometric maps (Van Eck et al., 2010). It allows the generation of various types of visualizations, including author, journal, and keyword, based on bibliometric data (Rusydiana et al., 2021). In the generated networks, the size of the nodes and labels reflects the significance of a document, while the thickness of the connecting lines between documents indicates the strength of their relationship, with thicker lines representing stronger links (Pascal & Renaud, 2020).

3.2. Data Selection

The Scopus database was utilized as the source for data collection in this study. Scopus was launched by Elsevier in 2004, and it contains over 15,000 peer-reviewed journals and offers broader coverage compared to Web of Science (Walsh & Renaud, 2017). Additionally, Scopus provides detailed information about each author in a cited reference (Zupic & Čater, 2015) and has been used as the database for bibliometric analysis (e.g. Donthu, Kumar, Pandey, et al., 2021; Pascal & Renaud, 2020).

Data were extracted from the Scopus database using the following search terms: (“design science” OR “design research”) AND “Information System”; (“design science” OR “design research”) AND (“behavioral science” OR “behavioral research”); and (“design science” OR “design research”) AND (“user perception” OR “user behavior” OR “user satisfaction” OR “user acceptance” OR “cognition” OR “user adoption” OR “user resistance” OR “organizational effectiveness” OR “Organization impact” OR “social relation”).

To maintain a comprehensive dataset, journals and conference papers were selected from Decision Support Systems (DSS), European Journal of Information Systems (EJIS), Information Systems Journal (ISJ), Information Systems Research (ISR), Journal of the Association for

Information Systems (JAIS), Journal of Information Technology (JIT), Journal of Strategic Information Systems (JSIS), Journal of Management Information Systems (JMIS), MIS Quarterly (MISQ), Information and Management (IM), The Americas Conference on Information Systems (AMCIS), the International Conference on Information Systems (ICIS), and the International Conference on Design Science Research in Information Systems and Technology (DESRIST). The search covered the period from 2004 to 2024, providing a 20-year understanding of how IS scholars are engaging with design science and behavioral science paradigms. Using the specified search terms, a total of 792 records were retrieved from the Scopus database. The data was cleaned by removing duplicates and excluding articles that were not relevant to design science research. Articles that employed the use of action design research were also excluded, along with books, editorial notes, and conference reviews. After this process, 528 papers remained. The results of the search are summarized in Table 1 below.

Search terms	Result
("design science" OR "design research") AND "information system"	733
("design science" OR "design research") AND ("behavioral science" OR "behavioral research")	44
("design science" OR "design research") AND "user perception" OR "user behavior" OR "user satisfaction" OR "user acceptance" OR "cognition" OR "user adoption" OR "user	15

resistance” OR “organizational effectiveness” OR “Organization impact” OR “social relation”)	
Total result	792

Table 1: Search Results from Scopus Database.

4. Results and Discussion

4.1. Analysis of Design Science Research in Information Systems

4.1.1. Trends of design science publication per year

We visualized the yearly trends in design science publications for the period from 2004 to 2024, as illustrated in Figure 1.

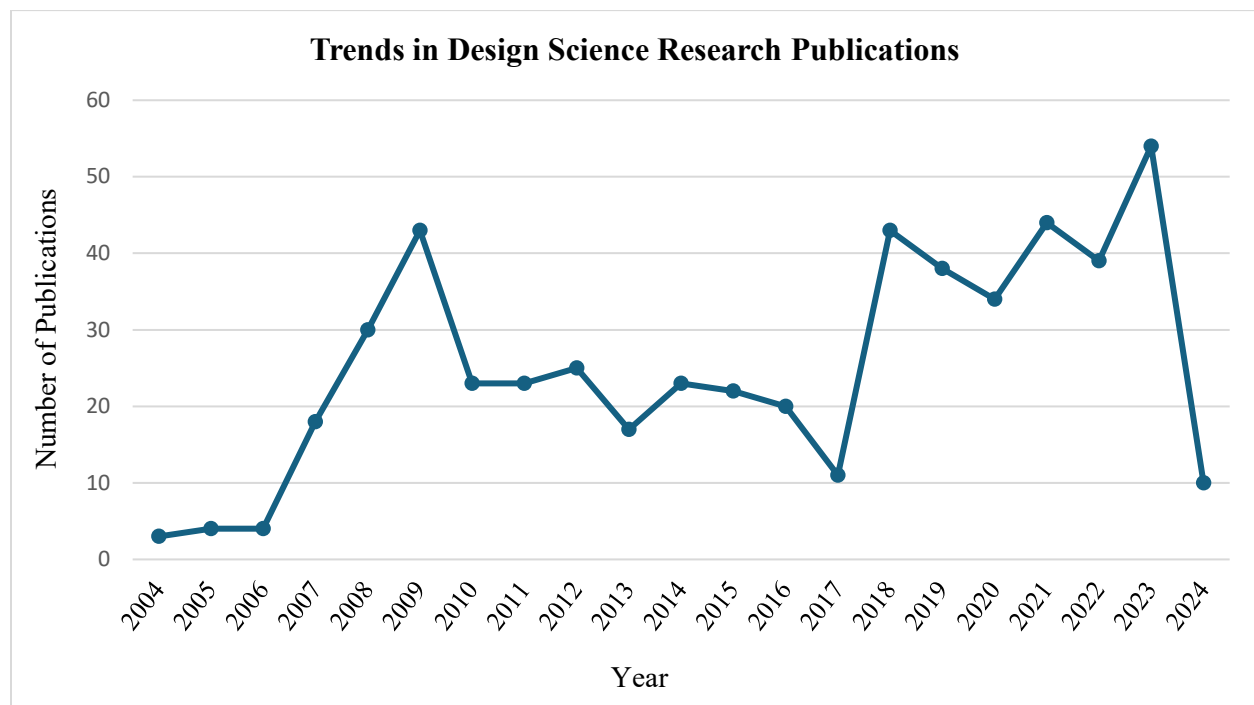


Figure1: Trends in Design Science Research Publications

As shown in Figure 1, the trend in design science research publication across the IS journal and conference outlets under review from 2004 to 2024 provides valuable insights into the engagement of IS scholars in this domain. The findings reveal significant growth and fluctuations in publication activity. Between 2004 and 2009, there was a steady increase in the publication, peaking in 2009, which may be attributed to heightened interest and foundational advancement of design science research in the field. A decline and stabilization follow between 2010 and 2017, suggesting a maturing research area with steady contributions. From 2018, renewed growth is observed, reaching a peak in 2023, highlighting renewed interest of IS scholars in design science research. However, 2024 shows a sharp decline in publications across the outlets under review. This could be attributed to shifts in research towards emerging areas such as AI-driven IS research. Overall, these findings highlight the growing interest of IS scholars in design science research and the dynamic publication pattern across the IS journal and conference outlets being reviewed.

4.1.2. Most Influential Publication

A citation analysis was performed on the data generated. Using a threshold of a minimum of 20 citations for each article. Out of the 528 articles on design science extracted from the dataset, 89 articles met the threshold of a minimum of 20 citations. The resulting citation analysis visualization highlights influential works and their relationships within the domain of design science research. The bigger the node and its label, the more influential the research paper is. The network visualization of the most influential publications is shown in Figure 2 below.

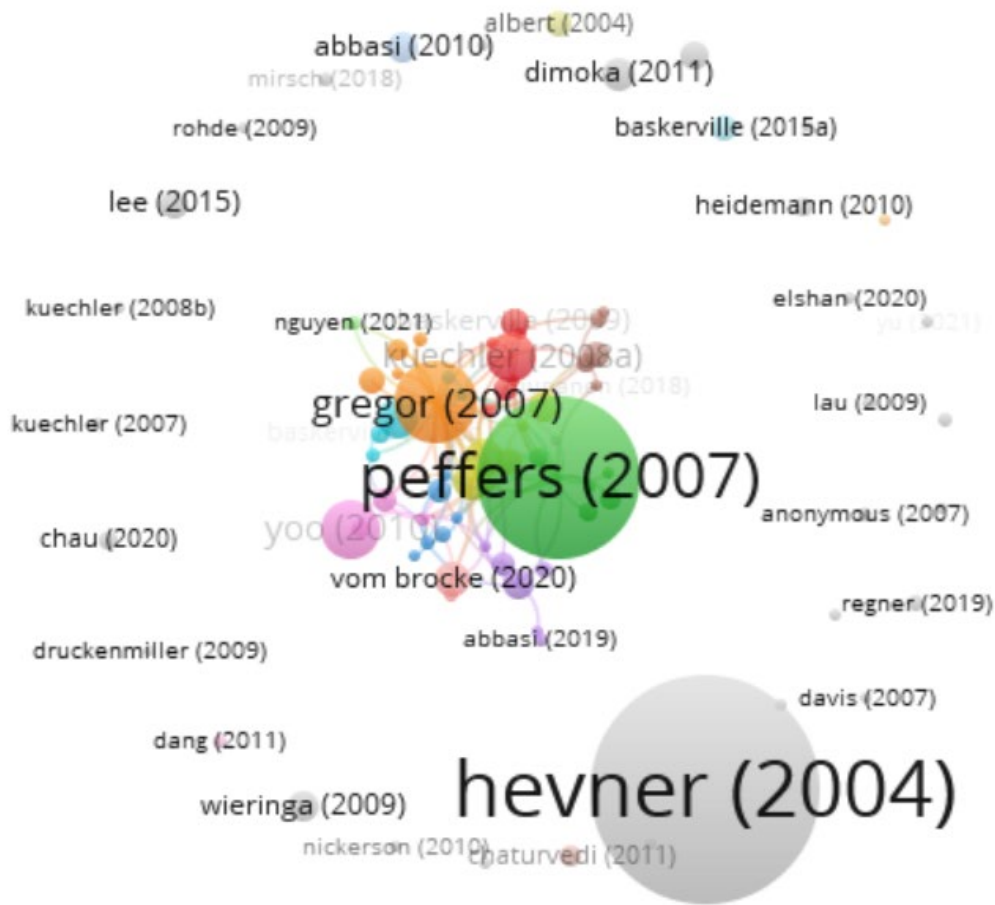


Figure 2: Citation Analysis of Design Science Publications

The citation analysis result presented in Figure 2 highlights some publications as pivotal reference points in the development of design science research. The seminal work by Hevner et al. (2004) emerges as the most influential publication, with a total of 10,050 citations. Following this, the seminal work by Peffers et al. (2007) ranks second with a total of 5,150 citations. Table 2 below presents a list of the top five influential publications and the number of citations extracted from the citation analysis result presented in Figure 2.

S/N	Author(s)	Article	Number of Citations
1	Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004)	Design Science in Information Systems Research.	10050
2	Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007)	A Design Science Research Methodology for Information Systems Research.	5150
3	Gregor, S., & Jones, D. (2007)	The Anatomy of a Design Theory.	1332
4	Yoo, Y. (2010)	Computing in everyday life: A call for research on experiential computing.	717
5	Kuechler, B., & Vaishnavi, V. (2008)	On theory development in design science research: anatomy of a research project.	509

Table 2: Top Five Influential Publications

4.1.3. Distribution of Design Science Research Publications Across Selected IS Journals and Conference Outlets.

The distribution of 528 articles across IS journal and conference publication outlets was visualized using Excel, as shown in Figure 3 below.

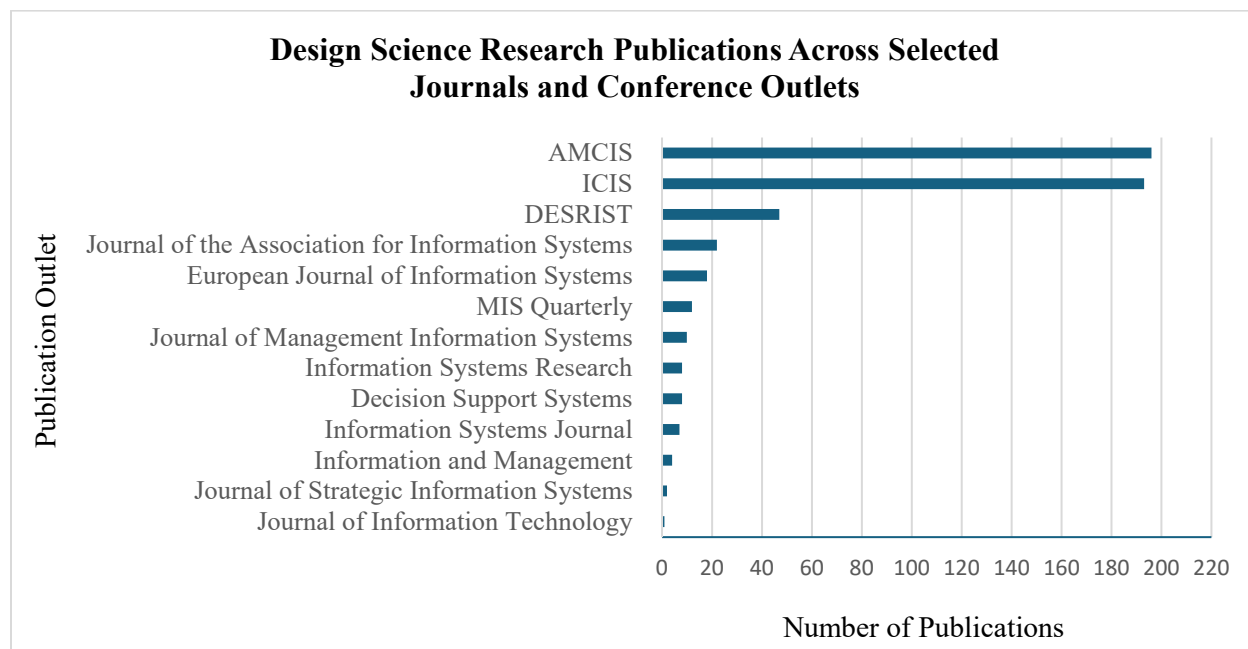


Figure 3: Design Science Research Publications Across Selected Journals and Conference Outlets

The analysis of design science research publications across selected IS journals and conference outlets highlight varying levels of contributions. AMCIS and ICIS emerge as the leading publication venues, with the highest number of publications. This indicates the central role of these two publication outlets in disseminating design science research. DESRIST follows as a significant contributor, reflecting its focus on advancing design science research. Journals such as the Journal of the Association for Information Systems and the European Journal of Information Systems contribute a moderate number of publications, emphasizing their engagement with the publication of design science research. Some journals, such as Information and Management, Journal of Strategic Information Systems, and the Journal of Information Technology, show a relatively lower representation of design science publications. Overall, the

results indicate the dominance of conference outlets in promoting design science research and suggest opportunities to expand design science presence in underrepresented journals.

4.1.4. Co-citation Analysis of Publications

Co-citation analysis was performed to understand the citation patterns and research focus of design science scholars in the IS field. Authors referencing the same documents are identified to have some commonalities in terms of research focus. Out of the 22281 cited references in our dataset, only 46 cited references met the threshold of a minimum of 7 citations. The clusters that emerged are presented in Figure 4 below.

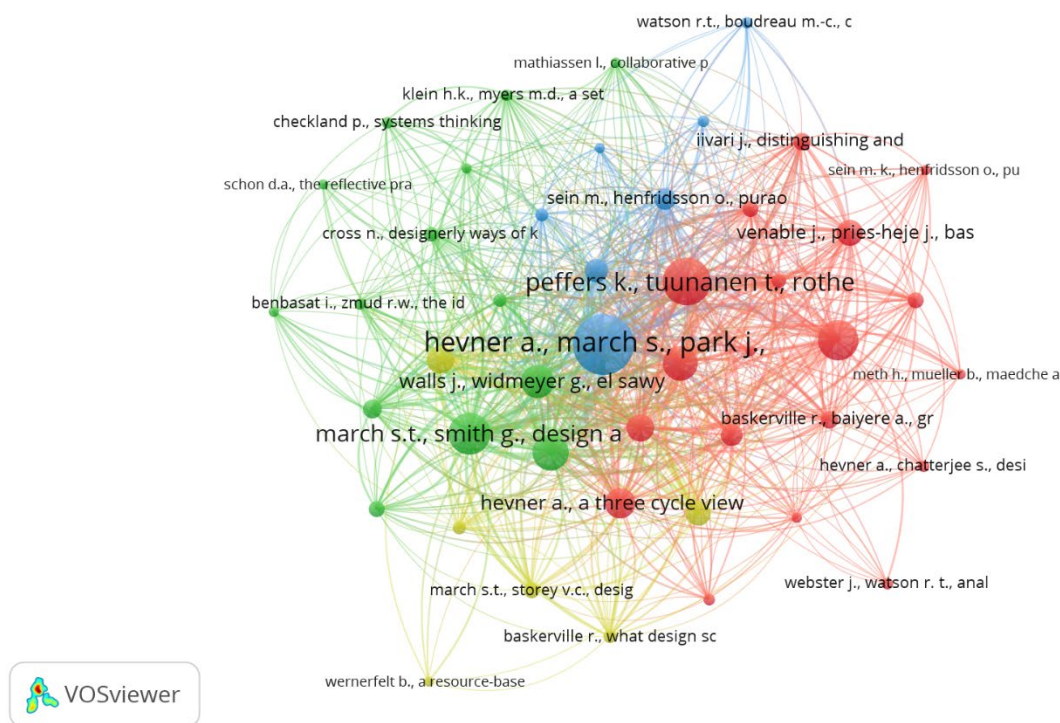


Figure 4: Co-citation Analysis of Publications

Based on the co-citation analysis results above, four clusters emerged. The key publications were identified based on the size of the node. These publications were used to identify the focus of each theme. The theme of the clusters that emerged are listed below:

1. Cluster 1 (Red): Design science research methodology.
2. Cluster 2 (Green): Understanding design science and natural (behavioral) science
3. Cluster 3 (Blue): Understanding IS research methods (Design Science and Action Design Research)
4. Cluster 4 (Yellow): Design science theory

The themes identified above suggest that most of the discussions in the outlets under review focus on guidelines for conducting design science research. In addition, they explore topics such as the synergy between design science and behavioral science, action design, and the development of design theory.

4.1.5. Co-Authorship Analysis

Co-authorship analysis is used to understand the intellectual interactions among researchers in a research field (Donthu, Kumar, Mukherjee, et al., 2021). 1180 authors contributing to design science research in the publication outlets under investigation were analyzed using co-authorship analysis to determine their research collaboration. The co-authorship analysis was performed using a threshold of a minimum of 20 citations to the author's work and the author should have published at least 3 articles. Authors with variations in their names were harmonized using the Thesaurus VOSviewer file. For instance, Hevner Alan's name appears as hevner alan r. and hevner alan. Similarly, Baskerville Richard's name appears as 183rexel183183lle, 183rexel183 l.

and 184rexel184184lle, Richard. Out of the 1180 authors, only 53 authors met these criteria. The result of the co-authorship analysis is presented in Figure 5 below.

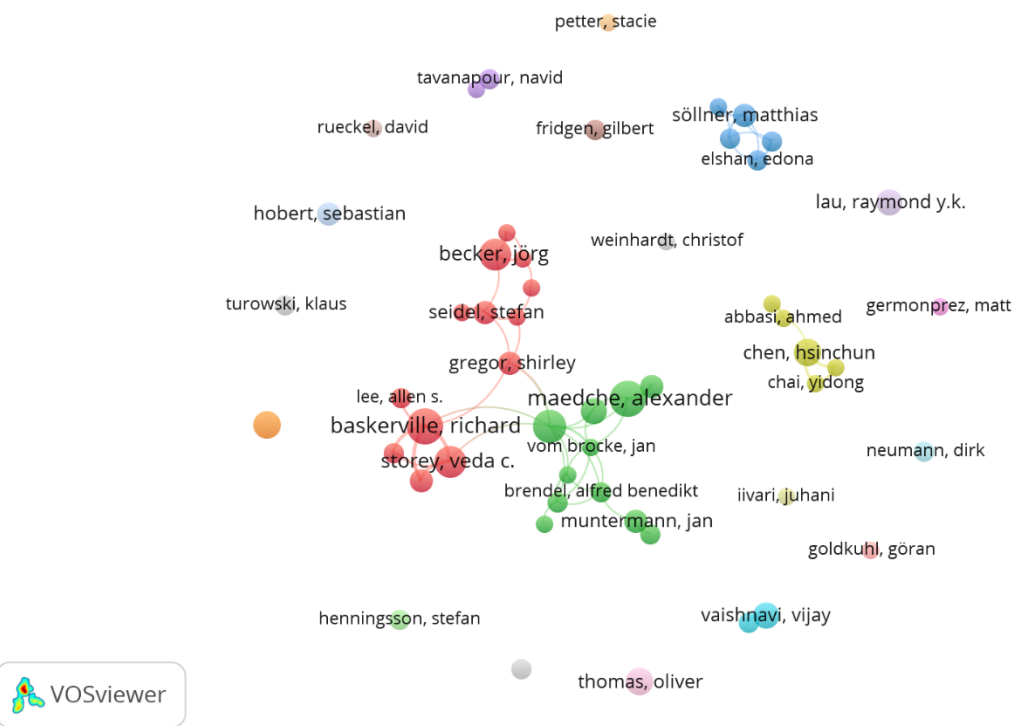


Figure 5: Co-authorship Analysis

The result of the co-authorship analysis presented in Figure 5 shows moderate collaboration among the 53 authors with only 24 authors engaging with between cluster collaboration. These findings indicate that collaboration is more prominent within clusters than there is between clusters. This suggests that design science scholars in the IS field interact mainly with other scholars within their research stream. Table 3 below shows the list of the top ten influential authors among the 53 authors who met the minimum criteria for selection.

S/N	Author	Number of Documents	Number of Citations

1	Hevner Alan	10	10368
2	Tuunanen Tuure	4	5224
3	Gregor Shirley	5	1792
4	Vaishnavi Vijay	6	847
5	Baskerville Richard	12	846
6	Kuechler Bill	4	570
7	Seidel Stefan	5	479
8	Maedche Alexander	12	439
9	Chandra Kruse, Leona	3	438
10	Pries-heje Jan	4	408

Table 3: List of the Top Ten Influential Authors in Design Science

4.2. Complementary use of Design Science and Behavioral Science Research in IS

To answer our second research question (RQ2) on how IS scholars have been engaging with the complementary use of design science and behavioral science in their research, we analyzed the abstracts of each article in our dataset to see how behavioral science research has been integrated into the design and evaluation of the IT artifacts. Specifically, we focused on the evaluation phase, guided by Hevner et al. (2004) information systems research framework and the extended design research methodology by Sutton et al. (2021), which outlines the complementary use of design science research and behavioral science research. To assess the extent to which IS scholars engage with the complementary use of the design science and behavioral science paradigm, we examined whether the evaluation extended beyond functionality to include the artifact's influence on behavioral outcomes, organizational processes, or decision-making. Figure 6 below shows the distribution of research types in selected publication outlets.

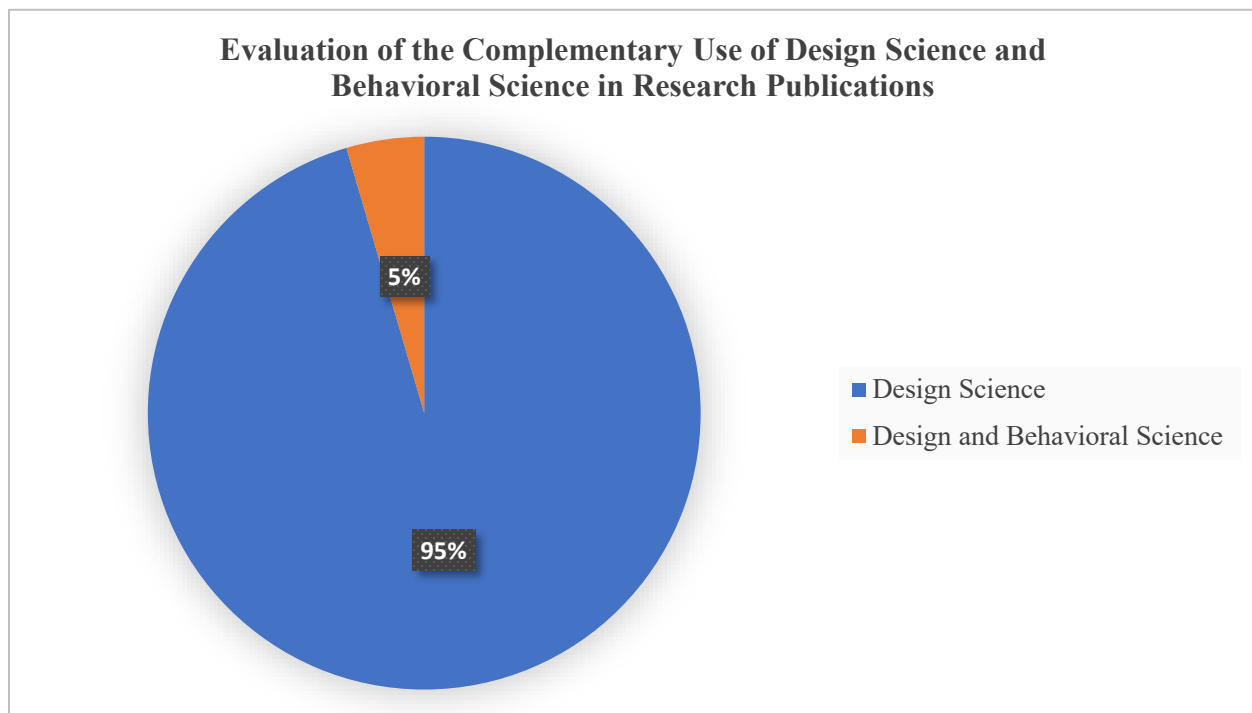


Figure 6: Evaluation of the Complementary Use of Design Science and Behavioral Science in Research Publications

We used Excel to illustrate the percentage of publications that focus on design science and the complementary use of the two paradigms. i.e. publications that mention the evaluation of the utility and impact of the IT artifacts on users and organizations. Our analysis reveals that 5% of the publications under review evaluate the utility and impact of their design artifact on users and organizations. This indicates that the integration of behavioral science research into design science research by Information Systems (IS) scholars remains limited. While utility testing is often emphasized, the impact of artifacts on user behavior and organizational dynamics is less frequently explored. This indicates a gap in leveraging the full potential of behavioral science

research to enhance design science research evaluations, highlighting an opportunity for IS researchers to adopt a more holistic approach in future studies.

5. Conclusion

Our findings underscore the dynamic nature of design science research within the IS field. While there has been considerable progress in publication volume, gaps remain in the complementary use of design science and behavioral science paradigms. The dominance of conference outlets suggests that design science research benefits significantly from collaborative and iterative discussion platforms, but the underrepresentation in certain journals signals the need for broader dissemination.

The limited engagement of IS scholars with behavioral science in design science evaluations highlights a significant gap in research practices. Although scholars in the IS field call for the complementary use of these paradigms, evaluations largely remain limited to assessing the utility of the IT artifacts. This lack of holistic evaluation limits the potential for understanding the impact of IT artifacts on user behavior and organizational success.

Future research can expand the scope of this study by examining more journals and conference publication outlets. In addition, other databases such as Web of Science and Google Scholar may be utilized to capture a broader range of publications. Also, a Delphi study can be conducted to reach consensus on the practicality of integrating the two paradigms in IS research.

Our research is not without limitations. It focuses mainly on a selection of some top IS journals and conference outlets. Also, the use of the Scopus database may have constrained the breadth of our search results. Furthermore, identifying the themes of clusters in the co-citation analysis is subjective, as it was based on the common topic shared by at least three most referenced articles

within each cluster. In addition, the analysis of the complementary use of the two paradigms relied solely on the abstracts of the articles.

References

- Antunes, P., Hoang Thuan, N., Johnstone, D., & Hoang, N. (2021). Nature and Purpose of Conceptual Frameworks in Design Science. In *Scandinavian Journal of Information Systems* (Vol. 33, Issue 2).
- Ayanso, A., Lertwachara, K., Lertwachara, K., & Vachon, F. (2011). Design and behavioral science research in premier IS journals: evidence from database management research. *International Conference on Design Science Research in Information Systems and Technology*. https://doi.org/10.1007/978-3-642-20633-7_10
- Ayanso, A., Lertwachara, K., & Vachon, F. (2007). Diversity or Identity Crisis? An Examination of Leading IS Journals. *Communications of The AIS*. <https://doi.org/10.17705/1cais.02042>
- Banville, C., & Landry, M. (1989). ARTICLES Social Aspects of Computing Can the Field of MIS be Disciplined? In *Communications of the ACM* (Vol. 32, Issue 1).
- Baskerville, R. (2008). What design science is not. In *European Journal of Information Systems* (Vol. 17, Issue 5, pp. 441–443). Palgrave Macmillan Ltd. <https://doi.org/10.1057/ejis.2008.45>
- Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design science research contributions: Finding a balance between artifact and theory. In *Journal of the Association for Information Systems* (Vol. 19, Issue 5, pp. 358–376). Association for Information Systems. <https://doi.org/10.17705/1jais.00495>
- Benbasat, I., & Weber, R. (1996). Research Commentary: Rethinking “Diversity” in Information Systems Research. In *Information Systems Research* (Vol. 7, Issue 4, pp. 389–399). INFORMS Inst. for Operations Res. and the Management Sciences. <https://doi.org/10.1287/isre.7.4.389>
- Culnan, M. J. (1986). The Intellectual Development of Management Information Systems, 1972–1982: A Co-Citation Analysis. *Management Science*, 32(2), 156–172. <https://doi.org/10.1287/mnsc.32.2.156>
- Deng, Q., & Ji, S. (2018). A Review of Design Science Research in Information Systems: Concept, Process, Outcome, and Evaluation. *Pacific Asia Journal of the Association for Information Systems*. <https://doi.org/10.17705/1pais.10101>
- Deng, Q., Wang, Y., & Ji, S. (2017). Design Science Research in Information Systems: A Systematic Literature Review 2001-2015. *International Conference on Information Resources*

Management.

<https://www.semanticscholar.org/paper/24554dd5f7a4127668411e5268ea424382789486>

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296.

<https://doi.org/10.1016/j.jbusres.2021.04.070>

Donthu, N., Kumar, S., Pandey, N., & Gupta, P. (2021). Forty years of the International Journal of Information Management: A bibliometric analysis. *International Journal of Information Management*, 57. <https://doi.org/10.1016/j.ijinfomgt.2020.102307>

Dwivedi, Y. K., Mustafee, N., Carter, L. D., Williams, M. D., & Bibliometric, A. (2010). *A Bibliometric Comparision of the Usage of Two Theories of IS/IT Acceptance (TAM and UTAUT)*.

<http://aisel.aisnet.org/amcis2010/183>

Fischer, C. (2011). The Information Systems Design Science Research Body Of Knowledge – A Citation Analysis In Recent Top-Journal Publications. *Pacific Asia Conference on Information Systems*. <https://www.semanticscholar.org/paper/af4becff38892dcd932b0ba904ea7f17ed0cfc5>

Goes, P. B. (2014). Editor's comments: design science research in top information systems journals. *Management Information Systems Quarterly*.

<https://www.semanticscholar.org/paper/8e1ef806fb624058fdc2078330ffb0949bccdae2>

Goldkuhl, G. (2016). Separation or unity? Behavioral science vs. design science. In *Proceedings of the 37th International Conference on Information Systems (ICIS) (Pp. 1-13)*., 1–13.

Hevner, A. R., March, S. T., Park, Jinsoo., & Ram, Sudha. (2004). Design science in information systems research. In *Source: MIS Quarterly* (Vol. 28, Issue 1).

Huysmans, P., & De Bruyn, P. (2013). *A mixed methods approach to combining behavioral and design research methods in information systems research*. http://aisel.aisnet.org/ecis2013_cr/29

Jain, H., Sinha, A. P., Vitharana, P., Ayanso, A., Lertwachara, K., & Vachon, F. (2011). Design and Behavioral Science Research in Premier IS Journals: Evidence from Database Management Research. In *LNCS* (Vol. 6629).

Kelton, A. S., & Murthy, U. S. (2023). Reimagining design science and behavioral science AIS research through a business activity lens. *International Journal of Accounting Information Systems*, 50. <https://doi.org/10.1016/j.accinf.2023.100623>

Lukyanenko, R. (2021). Design Behavioral Research: Toward a Unified View of Information Systems. *Americas Conference on Information Systems*.

<https://www.semanticscholar.org/paper/aa035bb51e9a6af71bbc98673bb2b8ebc831c7a2>

Maedche, A., Gregor, S., & Parsons, J. (2021). Mapping Design Contributions in Information Systems Research: The Design Research Activity Framework. *Communications of The Ais*.

<https://doi.org/10.17705/1cais.04914>

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. In *Decision Support Systems* (Vol. 15).

Metallo, C., Metallo, C., Metallo, C., Metallo, C., Agrifoglio, R., Agrifoglio, R., Agrifoglio, R., Agrifoglio, R., & Spagnoletti, P. (2012). *Design Science in Information Systems Research: foundational and evolutionary aspects*.

<https://www.semanticscholar.org/paper/cd8b7cc4a897fc33092715d53072a4f34ed9b06d>

Niederman, F., & March, S. T. (2012). Design science and the accumulation of knowledge in the information systems discipline. *TMIS*. <https://doi.org/10.1145/2151163.2151164>

Pascal, A., & Renaud, A. (2020). 15 Years of Information System Design Science Research: A Bibliographic Analysis. *Hawaii International Conference on System Sciences*.

<https://doi.org/10.24251/hicss.2020.618>

Rusydiana, A. S., Rahardjo, S., & Soeparno, W. S. (2021). Mapping Research on Halal Logistics using VoSviewer. *Library Philosophy and Practice*, 1–14.

<https://digitalcommons.unl.edu/libphilprac>

Sutton, S. G., Arnold, V., Collier, P., & Leech, S. A. (2021). Leveraging the synergies between design science and behavioral science research methods. *International Journal of Accounting Information Systems*, 43, 100536.

Thuan, N. H., Drechsler, A., & Antunes, P. (2019). Construction of design science research questions. *Communications of the Association for Information Systems*, 44(1), 332–363.

<https://doi.org/10.17705/1CAIS.04420>

Van Eck, N. J., Waltman, L., Dekker, R., & Van Den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405–2416.

<https://doi.org/10.1002/asi.21421>

- Vessey, I., Ramesh, V., Ramesh, V., & Glass, R. L. (2002). Research in Information Systems: An Empirical Study of Diversity in the Discipline and Its Journals. *Journal of Management Information Systems*. <https://doi.org/10.1080/07421222.2002.11045721>
- Walsh, I., & Renaud, A. (2017). Reviewing the literature in the IS field: Two bibliometric techniques to guide readings and help the interpretation of the literature. *Systèmes d'information & Management*, Volume 22(3), 75–115. <https://doi.org/10.3917/sim.173.0075>
- Wimmer, H., & Yoon, V. (2017). Counterfeit product detection: Bridging the gap between design science and behavioral science in information systems research. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2017.09.005>
- Zupic, I., & Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>

From Pixels to Insights: Large Language Models for Automated Data Visualization Enhancement

Dr. Andres Fortino
New York University
agf249@nyu.edu

ABSTRACT

This study explores the potential of Large Language Models (LLMs) as visualization coaches using a structured analytical framework. We trained ChatGPT with an 18-point visualization assessment framework and evaluated its ability to analyze and suggest improvements for data visualizations from a White House economic policy presentation. The research compared ChatGPT's analysis against a Subject Matter Expert (SME) across five diverse chart types, with agreement levels ranging from 81.4% to 93.3%. The highest concordance appeared in functional aspects of visualization (>90%), while the greatest divergence emerged in technical recommendations and chart type selection (81.4%). ChatGPT successfully generated improved versions of the charts, focusing on enhanced clarity and accessibility for general audiences. While the results demonstrate the feasibility of using AI as a visualization coach, they also highlight important limitations in technical depth and context awareness. This proof-of-concept study suggests that LLMs can effectively complement human expertise in data visualization, particularly for basic analysis and accessibility improvements, while indicating areas needing further development for more sophisticated applications.

Keywords: Data visualization, Large language models, ChatGPT, Visual analytics, Human-AI collaboration, Decision support

INTRODUCTION

Our study aims to develop a chatbot that analyzes and improves data visualizations. This tool will assist data analysts by automatically evaluating and suggesting enhancements to their charts, making them more effective for communication and decision-making. This project addresses the growing need for high-quality data visualizations in businesses to ensure that insights are accurately and effectively conveyed.

The Problem

In the current data-driven business environment, there is a growing need for high-quality data visualizations to communicate insights and support decision-making effectively. However, creating clear, accurate, and visually appealing data visualizations often requires specialized skills and expertise that not all data analysts possess. Traditional methods of improving data visualizations can be time-consuming and subjective, leading to inconsistent quality across different analysts and organizations. There is a need for a more efficient, standardized, and accessible approach to enhancing data visualizations that can assist data analysts in improving their charts consistently and effectively.

Traditionally, applying such a checklist would require significant time and expertise from a human analyst. However, the chapter also explores an exciting new frontier: the use of large language models (LLMs) in data visualization analysis. By fine-tuning an LLM with the principles embodied in the 18-point checklist, we create a powerful AI assistant capable of rapidly analyzing and suggesting improvements for data visualizations.

The paper concludes by comparing the manual analysis process with the AI-assisted approach. This comparison highlights both the progress made in automating complex analytical tasks and the continued importance of human judgment and creativity in data visualization. While the AI can quickly identify potential issues and suggest improvements based on established principles, the human analyst brings contextual understanding, creative problem-solving, and the ability to make nuanced decisions that go beyond rule-based assessments.

By presenting a comparison of both traditional and AI-assisted methods, this paper equips readers with a versatile toolkit for creating and improving data visualizations. Whether working through the list manually to develop a deep understanding of visualization principles or leveraging AI to streamline the analysis process, readers will gain valuable skills for communicating data effectively in various professional contexts.

Theory

The project is based on the concept that a chatbot utilizing Large Language Models (LLMs) and established data visualization principles can effectively analyze and offer valuable suggestions for improving data visualizations. This approach relies on the idea that AI can be trained to recognize and apply best practices in data visualization, as developed by Professor Fortino, to identify areas where charts may lack clarity, accuracy, and visual appeal. By automating this analysis and recommendation process, the chatbot could potentially provide consistent, objective, and actionable feedback for enhancing data visualizations, thereby improving the communication of insights across various professional settings.

Research Questions

1. How can an AI-driven chatbot effectively analyze and suggest improvements for data visualizations based on established principles?
2. To what extent can a chatbot provide valuable and actionable feedback for enhancing the clarity, accuracy, and visual appeal of data visualizations compared to traditional methods?
3. What is the potential impact of an AI-based visualization enhancement tool on the quality and consistency of data presentations in professional settings?

Hypothesis

An AI-driven chatbot, leveraging established data visualization principles, can provide valuable and actionable suggestions for improving data visualizations, resulting in measurable enhancements in clarity, accuracy, and visual appeal compared to the original charts, as demonstrated through expert evaluations and comparative analysis of visualizations before and after chatbot-suggested improvements.

LITERATURE REVIEW

The framework proposed by Fortino (2022) for analyzing and improving visuals consists of six dimensions: Story, Signs, Purpose, Perception, Method, and Charts. These dimensions provide a holistic approach to evaluating and enhancing data visuals. Each dimension focuses on specific aspects of effective visualization and is rooted in the work of known scholars in the discipline, as annotated below:

1. Story: Ensuring the visual tells a clear narrative
2. Signs: Using appropriate symbols and maintaining a high signal-to-noise ratio

3. Purpose: Aligning with organizational needs and audience expectations
4. Perception: Leveraging principles of visual perception for clarity
5. Method: Employing best practices in color use, decluttering, and titling
6. Charts: Selecting and implementing the most appropriate chart types

Within these dimensions, an 18-element checklist (three in each dimension) serves as a practical tool for assessing and refining data visualizations. It prompts creators to consider crucial elements such as the clarity of the visual story, the effectiveness of signs and symbols, the match between the visualization and its intended purpose, the application of perceptual principles, the judicious use of visual methods, and the selection of appropriate chart types.

Story

Al Shallowly (2011), an expert visual storyteller, tells us that: “Visualizations act as a campfire around which we gather to tell stories.” *Are we telling compelling business stories with our visuals?*

Create a Visual Story – Visual stories condense complex information into a clear, visual narrative that is easy to grasp. They should highlight the key message and be self-explanatory, minimizing the cognitive effort required to understand. As Alberto Cairo (2012) notes, effective visualizations align data representation with storytelling principles to captivate and inform. Key Question: *Does the visual make a clear point that would take many words to convey?*

Make It a Prop – A visualization becomes a prop when it facilitates engagement and drives meaningful conversations. It should encourage interaction, allowing stakeholders to explore insights collaboratively. Hans Rosling (2006) demonstrated this by using vivid graphics in presentations, turning static data into interactive and memorable experiences. Key Question: *Does the visual serve as a central tool for discussion and insight?*

Emulate Legendary Storytellers – Great visualizations emulate the techniques of master storytellers by being structured, engaging, and emotionally resonant. They guide the audience through a compelling narrative arc, ensuring clarity and impact. Edward Tufte (1983) emphasizes the importance of precision and aesthetics in visual storytelling to enhance understanding. Key Question: *Does the visual embody timeless principles of great storytelling?*

Signs

Jacques Bertin (2011) explains in *Semiology of Graphics* that: “Data is transformed into graphics to understand. A map, a diagram are documents to be interrogated. But understanding means integrating all of the data. In order to do this it’s necessary to reduce it to a small number of elementary data. This is the objective of the ‘data treatment’ be it graphic or mathematic.” We must become good sign makers and send strong unmistakable signals to our audience of the direction they should take in their decision-making.

Semiotics, Signs, and Signifiers – Semiotics, the study of signs, ensures that visuals use symbols and signs appropriately to convey meaning effectively. Signs must align with their intended meaning (signifier) and audience expectations. Jacques Bertin emphasizes the importance of avoiding cultural missteps or confusion to ensure clarity in data visuals (Bertin, 2011). Key Question: *Are signs and symbols used properly in the visual?*

Visualization as a Communication System – A good visualization acts as a communication system, where the message is encoded clearly, transmitted effectively, and decoded without confusion. The signal-to-noise ratio must be high to reduce distractions and ensure the audience grasps the intended meaning (Shron, 2014). Key Question: *Does the visual send a strong unmistakable signal?*

The Functional Art – Art vs. Clarity – Visuals should prioritize functionality over aesthetics. Clarity should never be sacrificed for artistic appeal. Alberto Cairo emphasizes that data visuals are tools to inform and support decisions, not artworks designed to entertain or evoke emotions (Cairo, 2012). Key Question: *Is the chart functionally informational rather than beautiful art?*

Purpose

Stephen Kosslyn (2006) writes in *Graph Design for the Eye and Mind*: “The usefulness of a graph can be evaluated only in the context of the type of data, the questions the designer wants the readers to answer, and the nature of the audience.” We must keep in mind who our audience is, what they want and need to know, and what key-framed analytical questions we answer to inform them.

Information Need – A visualization must directly address the information needs of both the requester and the organization. It should provide data critical to the mission, enabling decision-making and bridging knowledge gaps. Max Shron’s framework emphasizes aligning visuals with real organizational problems to ensure decisions are informed and needs are met (Shron, 2014). Key Question: *Does the chart fulfill organizational information needs?*

Audience – Account for the audience’s perspective, ensuring clarity and resonance. By doing so, the visual aids the audience in achieving their objectives without confusion or mismatch (Cairo, 2012). Key Question: *Does the chart allow for audience biases, needs, and journey?*

Answer Well-Framed Analytical Questions – Visuals should present key evidence answering well-framed analytical questions developed during analysis. These answers must be clear, actionable, and evident, ensuring the audience grasps the conclusions and their implications for action (Shron, 2014). Key Question: *Does the visual answer a well-framed analytical question?*

Perception

Henry David Thoreau (1884), in his *Journal*, observes that “Many an object is not seen, though it falls within the range of our visual ray, because it does not come within the range of our intellectual ray, i.e. we are not looking for it. So, in the largest sense, we find only the world we look for.” And his key admonition to us as data visualizers is: “The question is not what you look at, but what you see.” Looking happens in the eye, and seeing happens in the mind. *What are we helping our audience to see?*

Use the Eye-Brain System of Seeing – Visuals should guide the viewer’s eye to the most critical information unconsciously. Techniques such as Gestalt principles, color theory, and the removal of chart junk can direct attention effectively. Alberto Cairo and Yarbus emphasize designing visuals to leverage the natural shortcuts of the human eye-brain system (Cairo, 2012; Yarbus, 1967).

Key Question: *Does the eye of the viewer focus on the most important point being made?*

Employ the Gestalt Principles of Perception – Gestalt principles ensure the visual design emphasizes figure/ground differences, grouping, connectedness, and flow. These principles guide viewers intuitively to the chart’s main message, enhancing clarity and understanding. Effective use of Gestalt psychology helps create visuals that are aesthetically pleasing and informative (Cairo, 2012). Key Question: *Have the principles of the Gestalt psychology of perception been thoughtfully employed in the visual?*

Design with Quality – A quality visual relieves the tension of ignorance by effectively informing the audience. Christopher Alexander’s “quality without a name” concept highlights the importance of creating “alive” visuals that satisfy the viewer’s need for knowledge through effective design and pattern language (Alexander, 1979). Key Question: *Does the visual resolve the tension in the viewer between not knowing and knowing?*

Method

We must use color appropriately, make sure our charts are devoid of all visual noise, and make sure the PowerPoint slide has a title that gets to the point of your message.

Use Color Effectively – Color enhances visuals when used strategically. Sparing use, semantic correctness, and accessibility for color-blind viewers are essential. Key elements should use color to draw attention while avoiding overuse. Proper application ensures the viewer focuses on critical data without confusion or distraction (Tufte, 1983). Key Question: *Is color used judiciously and sparsely?*

Remove All Chart Junk – Removing unnecessary elements, or “chart junk,” sharpens focus and enhances comprehension. Visuals should balance the data-to-ink ratio and eliminate clutter like extraneous grid lines, labels, or colors. Richard Tufte emphasizes simplicity and clarity, ensuring every element contributes to the message (Tufte, 1983). Key Question: *Is the visual clear of unnecessary visual elements not leading to a clear point being made?*

Tell the Story with the Title – A clear, explicit title ensures viewers understand the chart’s purpose immediately. Titles should succinctly express the key business point while avoiding ambiguity. McKinsey’s method emphasizes direct titling as a storytelling tool to avoid misinterpretation and confusion (McKinsey & Company). Key Question: *Does the title of the chart convey the point being made with the chart?*

Charts

The Charts dimension emphasizes the importance of selecting appropriate chart types and table designs to enhance understanding. By leveraging Cleveland and McGill’s functional scale (1987), chart types can be matched to the level of accuracy required for judgment. The chosen chart must effectively answer the business question being posed. Tables, when used, should be readable and enhanced with elements like conditional formatting and Sparklines for emphasis. Visual clarity, functionality, and appropriateness are the cornerstones of effective charting in data visualization.

Use the Right Chart – Cleveland and McGill’s functional scale guides the selection of charts by matching the level of encoding to the accuracy required for comparisons. The right chart should align with the purpose of the visualization, the level of judgment needed, and the type of

question being answered (Cleveland & McGill, 1987). Key Question: *Does the type of chart being used match the level of judgment required?*

Select the Chart Type Effectively – Different chart types are suited for specific purposes, such as Pareto charts for contributions or pie charts for parts of a whole. The selected chart must address the intended business question and communicate insights clearly and effectively (Nussbaumer-Knaflic, 2015). Key Question: *Does the chart type used match the business question being answered?*

Enhance Table Data for Emphasis – Tables should prioritize readability and emphasis through effective use of white space, shading, and conditional formatting. Visual enhancements like Sparklines add clarity and insight, transforming tables into powerful tools for communication (Nussbaumer-Knaflic, 2015). Key Question: *Are referenceable visuals (tables) readable with appropriate conditional formatting and thumbnail graphs used for emphasis?*

Dimension	Element	Key Question
Story	Create a Visual Story	Does the visual make a clear point that would take many words to convey?
	Make It a Prop	Does the visual serve as a central tool for discussion and insight?
	Emulate Legendary Storytellers	Does the visual embody timeless principles of great storytelling?
Signs	Bertain: Semiotics, Signs, and Signifiers	Are signs and symbols used properly in the visual?
	Visualization as a Communication System	Does the visual send a strong unmistakable signal?
	Cairo: The Functional Art - Art vs. Clarity	Is the chart functionally informational rather than beautiful art?
Purpose	Information Need	Does the chart fulfill organizational information needs?
	Audience	Does the chart allow for audience biases, needs, and journey?
	Answer Well-Framed Analytical Questions	Does the visual answer a well-framed analytical question?
Perception	Use the Eye-Brain System of Seeing	Does the eye of the viewer focus on the most important point being made?
	Gestalt Principles of Perception	Have the principles of the Gestalt psychology of perception been thoughtfully employed in the visual?
	Design with Quality	Does the visual resolve the tension in the viewer between not knowing and knowing?
Method	Use Color Effectively	Is color used judiciously and sparsely?
	Remove All Chart Junk	Is the visual clear of unnecessary visual elements not leading to a clear point being made?
	Tell the Story with the Title	Does the title of the chart convey the point being made with the chart?
Charts	Right Chart	Does the type of chart being used match the level of judgment required?
	Selection – The Basic Charts	Does the chart type used match the business question being answered?
	Tables - Analyze or Tell?	Are referenceable visuals (tables) readable with appropriate conditional formatting and thumbnail graphs used for emphasis?

Figure 1 – Summary of the six dimensions and 18 elements with the key question to be answered by the data visualizer.

METHOD

This research explored the use of Large Language Models (LLMs) for data visualization analysis using a structured framework. The process consisted of four key phases:

Phase 1 – Framework Training: An LLM (ChatGPT) was fine-tuned using the Fortino Framework document, which outlines 6 dimensions and 18 points for evaluating data visualizations. The framework emphasizes story, signs, purpose, perception, method, and charts as key analytical dimensions.

Phase 2 – LLM Analysis: The trained LLM analyzed five different data visualization cases using the framework. For each case, the LLM provided detailed analysis across all 18 points, including strengths, shortcomings, and suggested improvements. The LLM maintained consistent structure in its analysis, evaluating each dimension systematically.

Phase 3 – SME Validation: A Subject Matter Expert (SME) independently analyzed the same five visualizations using identical framework parameters. The SME analysis served as a benchmark for evaluating the LLM's performance. The SME followed the same structured approach but drew from deep domain expertise in data visualization.

Phase 4 – Comparative Analysis: A separate LLM (Claude) performed detailed comparative analysis between the ChatGPT and SME evaluations. This analysis quantified agreement levels for each framework point using percentage concordance and identified specific areas of alignment and divergence. The comparison revealed overall agreement ranging from 81.4% to 93.3%, with the strongest alignment in functional aspects and the greatest divergence in technical recommendations.

This methodology demonstrated the potential for LLMs to perform structured visualization analysis while highlighting areas where human expertise remains valuable. The high concordance levels suggest LLMs can effectively apply established frameworks, though they may emphasize different aspects than human experts.

The Prompt

You are a data visualization expert tasked with analyzing and improving charts using the principles outlined in the framework found in the “Data Visualization for Business Decisions” book I just uploaded. Your goal is to provide a comprehensive analysis of a given chart across all 6 dimensions and 18 analysis elements, each with specific questions to consider. Here’s your process:

- 1. Review the principles in the framework thoroughly, focusing on the 18 dimensions and their associated questions.*
- 2. When presented with a chart, carefully examine it in the context of each dimension.*
- 3. For each dimension: a. Answer all relevant questions from the checklist. B. Identify what aspects of the chart work well in this dimension. C. Pinpoint areas where the chart falls short in this dimension. D. Propose specific improvements to address these shortcomings.*
- 4. After analyzing all 6 dimensions and 18 elements, provide a summary of the key improvements needed, organized by dimension and element.*
- 5. Remember to focus solely on the principles and techniques outlined in the framework of the book, avoiding reference to external sources or methodologies.*
- 6. Maintain a critical yet constructive tone throughout your analysis.*

7. Be specific and detailed in your suggestions for improvement, providing clear rationale based on the book's principles.

8. Consider the chart in the context of its intended purpose and audience, as outlined in the framework.

Are you ready to analyze a chart and give me your opinion?

The Case

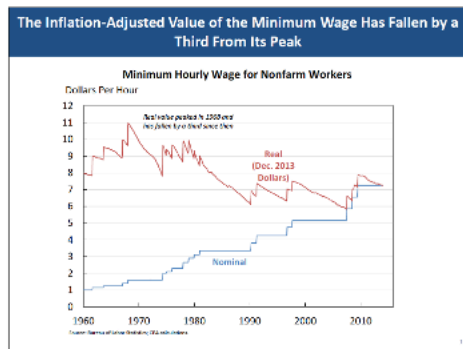
The case study is based on set of slides posted by the White House in the United States promoting a certain policy initiative. Early in President Obama's administration there was a push to have all states in the union as well as the Federal Government adopt a higher minimum wage level than was prevalent at that time. One of the agencies of the executive branch is the Council of Economic Advisors (2014). They laid out the economic case for raising the minimum wage. This slide set was made available at the time via the White House website to educate the public, policy makers, and lawmakers. We want to see if we can improve the charts in the presentation which was the chosen communication platform.

Think of yourself as a staff member of the Council of Economic Advisors, the group that created this presentation. One day you are called into the office of Chief Economist and he (or she) says: "Since you are one of our best data visualization experts on staff I would like to look over this document and tell us how we can improve these charts. We want to have the greatest impact for our readers." That's the basis of our analysis.

In this study, we shall undertake the analysis of additional slides from the case study. We have covered two of them in the lab book (slides 2 and 4 from the case study slide set). You can use the four additional examples here to further develop your understanding of how to improve data visual in a business context. To show you the usefulness of the analysis template and the process we will use it as the basis for analysis. You can use the template for yourself with each slide and then test your understanding against the remarks by the expert. Take this an assignment: "If I were given these slides to improve, what would I do along the lines of the six dimensions in the lab book, what would I change?" This way, you are further exercising your newly acquired visualization muscles to create compelling visuals.

The premise for our changes will be that there's a desire to reach and convince a more significant portion of the general population about the benefits of the proposed policy change. As it is, the case study charts were probably created to convince economists and policymakers and other trained people of the proposed policy changes. Although some in the general public could understand the charts, they are perhaps out of the reach of the average citizen. That's where you come in. There is a desire by management to create a new set based on these slides that are more accessible to people who not highly educated. Assume that the new slide set will be designed to be used in a presentation, not as a standalone document. Thus, they are to be a set of props for a presenter rather than a report to be read. Your mission: modify the current slides to fit the new audience and the new purpose.

Case 1 – Minimum Wage Original



Case 1 – Minimum Wage Revised

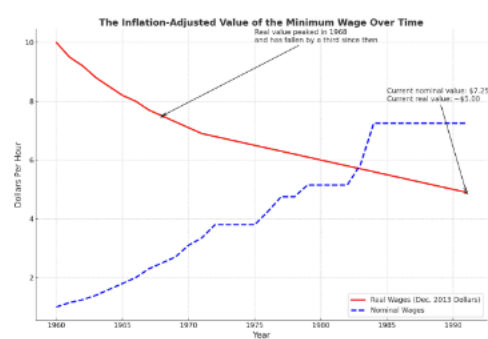


Figure 2 – Case 1 – Minimum Wage. Original chart and ChatGPT revised chart from the White House Economic Advisors Case dataset.

Example Outcome from ChatGPT

The analysis of Case 1 (minimum wage visualization) demonstrates ChatGPT's systematic application of the 6-dimension framework. ChatGPT's detailed assessment identified key strengths (clear trend visualization, appropriate chart type) and weaknesses (poor annotation visibility, unclear terminology for general audiences). The chatbot provided specific, actionable improvements focusing on accessibility and clarity, such as simplifying the title, enhancing line visibility, and adding explanatory notes for terms like "real" versus "nominal" wages. The revised visualization (Figure 2) shows these implemented changes, featuring clearer annotations, stronger color contrast, and simplified messaging. This analysis showcases both ChatGPT's ability to systematically evaluate visualizations and its practical capacity to generate improved versions, though some technical nuances noted by the SME were missed in the AI analysis.

ChatGPT's output advice to improve the visual:

"Story

"Visual Story: Strengths: The chart successfully conveys a clear story about the decline in the real value of the minimum wage since its peak in 1968. Shortcomings: The story could be stronger if the decline's implications for workers were explicitly linked to real-world outcomes. Improvements: Add a brief annotation or subtitle explaining the impact of the declining real value on workers' purchasing power."

"Visual Props: Strengths: The chart uses lines to compare real and nominal wages effectively. Shortcomings: Text annotations on the chart are somewhat small and may not grab the viewer's attention immediately. Improvements: Enlarge and bold the annotations to make them more visible and emphasize the key message."

"Storytellers: Strengths: The chart aligns with historical data storytelling techniques (e.g., timelines and comparisons). Shortcomings: Does not directly emulate iconic visuals for economic trends (e.g., Hans Rosling's engaging style or Florence Nightingale's clarity)."

Improvements: Incorporate elements such as callouts or a more engaging color scheme to make the visual more dynamic.

“Signs

“Signs: Strengths: Symbols (lines and labels) are used effectively to distinguish real and nominal values. Shortcomings: The red line for real wages is clear, but the blue line for nominal wages blends into the background. Improvements: Use a darker shade of blue or add a stronger outline to improve visibility.

Communication: Strengths: The signal-to-noise ratio is reasonably high as the chart focuses on its central message. Shortcomings: The title is slightly long and could distract from the chart. Improvements: Simplify the title to something like “Real Minimum Wage Has Dropped by One-Third Since 1968.”

Function: Strengths: The chart prioritizes clarity over aesthetic appeal, which is appropriate for the audience. Shortcomings: The design is functional but could be more visually engaging without sacrificing clarity. Shortcomings: Add subtle visual effects, such as shading or arrows, to emphasize key points.

“Purpose

“Need - Strengths: The chart fulfills its purpose of showing the historical decline in real wages. Shortcomings: It doesn’t clearly address why this decline is important or what action should follow. Improvements: Add a concluding sentence or callout to connect the data to the advocacy for raising the minimum wage.

“Audience – Strengths: The chart is designed for policymakers and an informed audience. Shortcomings: The general public may struggle with interpreting “real” versus “nominal” values without a brief explanation. Improvements: Include a simple legend or brief note defining “real” and “nominal.”

Frame - Strengths: The chart answers a well-framed analytical question about the historical trend of minimum wages. Shortcomings: The framing could be stronger if it directly tied the data to actionable insights. Improvements: Frame the title or subtitle as a question, e.g., “Why Has the Real Minimum Wage Declined by One-Third Since 1968?”

“Perception

“Seeing - Strengths: The red line for real wages naturally draws the eye, emphasizing the most critical trend. Shortcomings: The gridlines and axis labels are visually prominent, potentially distracting from the data. Improvements: Reduce the intensity of gridlines and make axis labels more subtle.

“Mind – Strengths: Gestalt principles, such as continuity and grouping, are evident in the use of lines. Shortcomings: The figure-ground distinction could be improved; the chart lacks a clear emphasis on the starting and ending points of the trend. Improvements: Highlight the peak and trough points with markers or annotations to improve focus.

“Quality – Strengths: The chart informs the viewer and dispels some ignorance about historical wage trends. Shortcomings: It doesn’t fully resolve the tension of understanding how these trends affect workers today. Improvements: Include a brief note or callout explaining the current purchasing power of minimum wages in 2013 dollars.

“Method

“Color – Strengths: Red and blue are used to differentiate real and nominal wages. Shortcomings: The blue line’s lighter color makes it less impactful. Improvements: Use a darker blue or bold the line for nominal wages.

“Chart Junk – Strengths: The chart avoids unnecessary elements like decorative icons or excessive annotations. Shortcomings: The text annotations could be perceived as clutter because they are small and overlap the lines. Improvements: Relocate annotations to the margin or use callout boxes.

“Title – Strengths: The title captures the essence of the chart’s message. Shortcomings: It’s long and somewhat redundant given the data labels on the chart. Improvements: Simplify the title to be more direct and impactful.

“Charts

“Right Chart – Strengths: The line chart is an appropriate choice for showing trends over time. Shortcomings: It might not convey the dramatic nature of the decline as effectively as an area or bar chart could. Improvements: Consider an area chart to emphasize the magnitude of the decline in real wages.

“Selection – Strengths: The chart type matches the question of wage trends over time. Shortcomings: It could better highlight key milestones, such as policy changes affecting wages. Improvements: Add vertical markers or shaded regions to indicate key historical events or policies.

“Tables – Strengths: Not applicable here as there are no tables in the visual. Shortcomings: None.

“Summary of Key Improvements

‘Title: Simplify to emphasize the key message more directly.

“Annotations: Enlarge or relocate text annotations for better visibility and focus.

“Color: Adjust the blue line to make it more distinct.

“Emphasis: Highlight key points on the lines (e.g., peaks and troughs) with markers or arrows.

“Audience Clarity: Add definitions or notes for “real” and “nominal” to aid comprehension.

“Context: Connect the trends to present-day implications or actions.”

RESULTS

The comparison between ChatGPT and SME analyses using the Fortino Framework reveals strong overall alignment, with concurrence ranging from 81.4% to 93.3%. The analyses show the strongest agreement in evaluating fundamental visualization principles, particularly in

function over decoration, basic chart type selection, and table analysis considerations. The SME consistently provides more historically grounded and technical recommendations, often referencing visualization pioneers like Playfair and emphasizing consistency with existing presentation sets. In contrast, ChatGPT tends to focus on modern design principles, accessibility considerations, and alternative visualization methods.

Five charts were selected from the White Hour Advisor's case that are representative of the most basic data visualization types: line, bar, pie, and time series. The first chart can be seen in Figure 1. The other four charts from the case are shown in Figure 2.

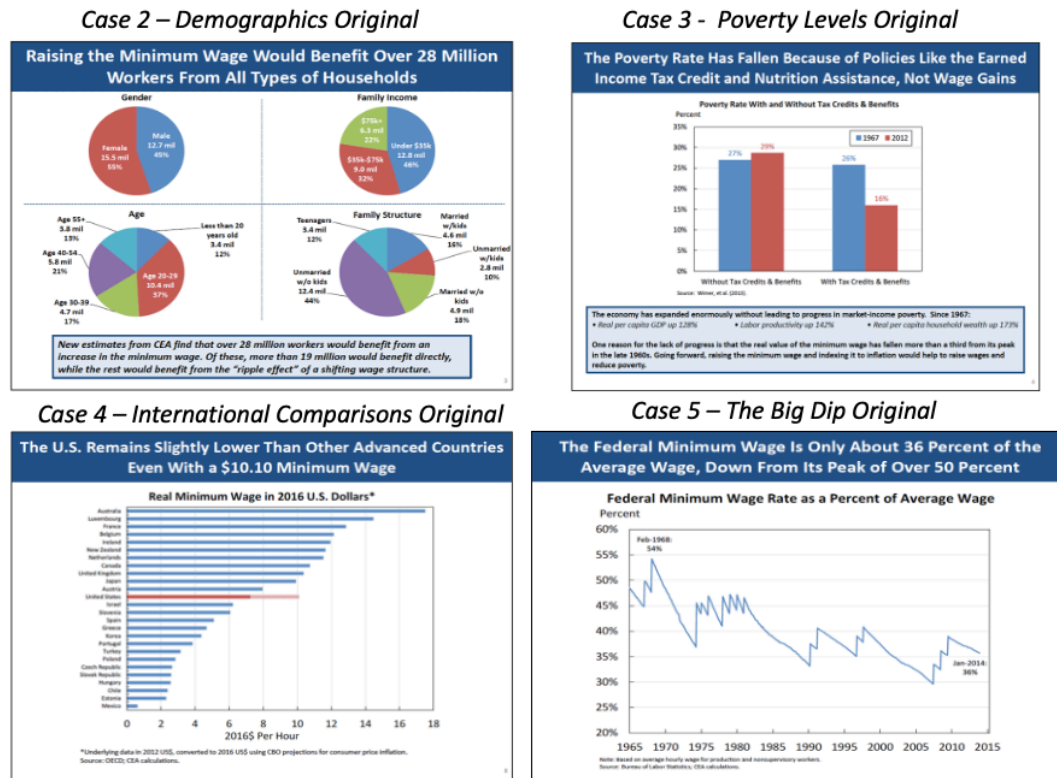


Figure 3 – Original Charts from the White House Economic Advisors Case dataset.

Key disagreements emerge in complex chart type recommendations, where the SME often favors traditional approaches while ChatGPT suggests modern alternatives. Notable differences also appear in color scheme recommendations, with ChatGPT emphasizing accessibility for colorblind viewers while the SME focuses on maintaining existing presentation schemes. The highest concurrence (93.3%) appears in Case 4's international comparison analysis, where both assessors strongly agreed on the effectiveness of the visualization, while the lowest agreement (81.4%) occurs in Case 2's demographic analysis, primarily due to divergent views on pie chart usage versus alternative formats.

Case Study	Overall Agreement	Strongest Agreement Areas	Major Disagreements
Case 1 (Minimum Wage)	85.5%	Function, Frame, Title	Chart type recommendations, Technical depth
Case 2 (Demographics)	81.4%	Function, Tables, Purpose	Chart type (pie vs. bar), Color usage
Case 3 (Poverty Levels)	84.2%	Function, Tables, Title	Chart type (line vs. bar), Visual emphasis
Case 4 (International Comparison)	93.3%	Function, Tables, Story	Color accessibility, Data grouping
Case 5 (Big Dip)	88.3%	Function, Tables, Communication	Historical context, Technical specificity

Figure 4 – Results of comparing SME analysis to ChatGPT analysis based on the 6-dimension framework for five separate visuals from the case.

The resulting reworked charts by ChatGPT are shown in Figure 5 for Cases 2 to 4. The reworked chart result for Case 1 is displayed in Figure 1. These reworked charts all show several characteristics: (1) they were generated by ChatGPT using Python code, (2) they appear to be an improvement on the original and useful for fulfilling the request to make them more understandable to the general public, and (3) they could use additional work to enhance them more compelling.

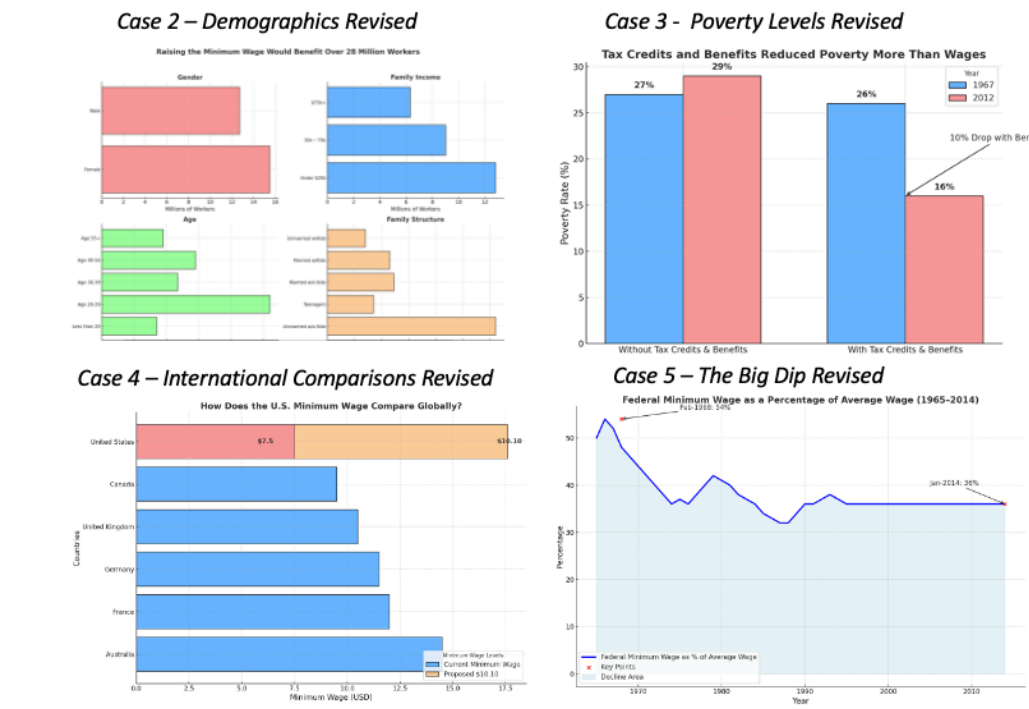


Figure 5 – ChatGPT revised charts from the White House Economic Advisors Case dataset.

DISCUSSION

The results of this exploratory study suggest that an AI-driven chatbot can meaningfully contribute to data visualization analysis and improvement, though with important caveats and limitations. The relatively high agreement rates (81-93%) between chatbot and expert analyses indicate promising potential, particularly in basic functional assessments and accessibility improvements. However, these findings should be interpreted as preliminary evidence rather than definitive proof of effectiveness. The chatbot's occasional divergence from expert recommendations, especially in technical aspects and traditional visualization approaches, highlights areas needing refinement. While the improved charts demonstrate enhanced clarity for general audiences, this proof-of-concept study primarily establishes the feasibility of using AI as a visualization coach rather than validating it as a replacement for human expertise. Further research is needed to rigorously validate these initial findings, particularly through controlled studies with larger samples and diverse visualization types. The results suggest this approach merits additional investigation while acknowledging that the current implementation represents just a first step toward developing reliable AI-assisted visualization tools.

CONCLUSIONS AND FURTHER WORK

This exploratory study demonstrated the feasibility of using LLMs as visualization coaches using a structured analytical framework. Our initial results, showing 81-93% agreement with expert analysis, suggest that AI can meaningfully assist in improving data visualizations while highlighting areas needing refinement.

Future work will focus on (1) Evaluating alternative frameworks, including Stephen Few's methodology (2012) and ChatGPT's native visualization knowledge; (2) Implementing Batova's 5-second and task-based testing protocols for more rigorous validation; (3) Conducting controlled usability studies comparing original and AI-improved visualizations; (4) Expanding the test dataset to include diverse visualization types and complexity levels; and (5) Investigating the role of domain expertise in AI-assisted visualization enhancement.

A key priority will be validating whether AI-improved visualizations consistently enhance understanding and decision-making across different user groups and contexts. This will include measuring both objective performance metrics and subjective user experience factors. We also aim to explore how AI coaching can be effectively integrated into existing visualization workflows while maintaining appropriate human oversight and creative control. This includes developing guidelines for when and how to apply AI suggestions most effectively.

REFERENCES

- Alexander, C. (1979), *The timeless way of building*. Vol. 1. New York: Oxford University Press, 1979.
- Argyris, C., and Schon, D. (1974). *Theory in practice: Increasing professional effectiveness*. Jossey-Bass, 1974.
- Batova, T. (2024). *Usability testing for data visualizations* (UVA-BC-0301). Darden School Foundation, University of Virginia.
- Bertin, J. (2011), "General theory, from semiology of graphics." *The Map Reader: Theories of Mapping Practice and Cartographic Representation* (2011): 8-16.

- Cairo, A. (2012), *The Functional Art: An introduction to information graphics and visualization*. New Riders, 2012.
- Cleveland, W., and McGill, R. (1987), "Graphical perception: The visual decoding of quantitative information on graphical displays of data." *Journal of the Royal Statistical Society. Series A (General)* (1987): 192-229.
- Cotgreave, A. (2013), "The 5 Most Influential Data Visualizations of All Time", Tableau, 2013, <https://www.tableau.com/learn/webinars/5-most-influential-visualizations-all-time>.
- Council of Economic Advisors. (2014). *The economic case for raising the minimum wage*. The White House. <https://www.slideshare.net/slideshow/the-economic-case-for-raising-the-minimum-wage/31128477>
- Few, S. (2009), *Now you see it: simple visualization techniques for quantitative analysis*. Analytics Press, 2009.
- Few, S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten* (2nd ed.). Analytics Press, 2012.
- Fortino, A. (2020). *Data Visualization for Business Decisions: A Laboratory Manual*. Mercury Publishing, 2020.
- Kosslyn, S. (2006), *Graph design for the eye and mind*. OUP USA, 2006.
- Lewis, R. (2010). *When cultures collide*. Nicholas Brealey Publishing, 2010.
- Medina, J. (2012). *Brain Rules*. Pear Press, Seattle, 2012.
- Nussbaumer Knaflic, C. (2015), "Story Telling with Data: Visualizing Opportunity", in September 16, 2015, Makeovers, <http://www.storytellingwithdata.com/blog/2015/9/16/visualizing-opportunity>.
- Rosling, H. (2006), "Hans Rosling shows the best stats you've ever seen, 2006." *Internet*: [http://www. Ted. Com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.Html](http://www.Ted.Com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.Html).
- Shalloway, A. (2011). Visualizations act as a campfire around which we gather to tell stories. *Net Objectives*. Retrieved from <https://visualign.org/quotes/>
- Shron, M. (2014), *Thinking with Data: How to Turn Information into Insights*. O'Reilly Media, Inc., 2014.
- Thoreau, H. (1884), *The Writings of Henry David Thoreau (Volume V)*. (1884).
- Tufte, E., and Graves-Morris, P. (2014), "The visual display of quantitative information; 1983." (2014).
- Yarbus, A. (1967), "Eye movements during perception of complex objects." *Eye movements and vision*, pp. 171-211. Springer, Boston, MA, 1967.

Towards a Data Science Pattern Language

Tianyu Wang
New York University
TW2798@nyu.edu

Andres Fortino
New York University
agfortino@nyu.edu

Roy Lowrance
New York University
roy.lowrance@gmail.com

ABSTRACT

This study introduces a data science pattern language to address inefficiencies and ethical challenges in data science workflows. The work stems from a recognized need for standardization in a field often plagued by ad hoc methodologies, lack of reproducibility, and ethical concerns such as bias and transparency. The project employed a two-step process: first, discovering reusable patterns that encapsulate best practices for common data science tasks such as data cleaning, feature engineering, and model evaluation, and second, developing an AI-driven chatbot to apply these patterns in real time. The pattern library was structured using the Alexandrian form, offering clear, actionable solutions, and was validated through expert feedback and application to real-world problems. The chatbot was designed to provide seamless access to these patterns, tested extensively for accuracy and usability, and evaluated for its ability to enhance decision-making processes. An exploratory validation comparing the chatbot's pattern classifications to a subject matter expert's assessments demonstrated fair agreement, indicating the chatbot's potential as a decision-support tool in data science workflows. Validation efforts, including a case study on predictive modeling in a retail context, demonstrated the framework's utility in improving workflow efficiency and ethical compliance. By integrating standardized solutions with real-time guidance, this project offers tools to streamline data science processes while fostering accountability and reproducibility.

KEYWORDS: Data science pattern language, Alexandrian pattern language, AI-guided workflow automation, Real-time decision support tool

INTRODUCTION

The rapid growth and adoption of data science across industries have revolutionized decision-making processes. This growth makes organizations derive insights from vast datasets. However, despite its transformative potential, data science workflows are often plagued by inefficiencies, inconsistencies, and ethical challenges. These issues arise from ad hoc methodologies, fragmented processes, and a lack of standardized practices, which hinder productivity but could compromise the reproducibility and scalability of data science projects (Egger et al., 2022).

The creation of a data science pattern language was undertaken to address these challenges. The project's primary objective was to develop a reusable library of best practices that standardize solutions for common data science problems. These patterns were complemented by an AI-driven chatbot designed to provide real-time guidance using the patterns, which helps data scientists access and apply these patterns efficiently within their workflows. By integrating ethical principles such as fairness, transparency, and bias mitigation, this project also bridges the gap between technical innovation and responsible AI deployment.

This paper discusses the development process, findings, and implications of the project. It shows how the pattern library and chatbot promote efficiency, reproducibility, and accountability in data science workflows and how they become valuable tools for both practitioners and organizations. Moreover, the integration of ethical considerations into the patterns ensures that the solutions align with societal expectations and regulatory requirements, particularly in sensitive domains such as healthcare and finance (Egger et al., 2022).

The creation of a data science pattern language exemplifies the intersection of innovation and ethics. It offers a scalable framework that not only addresses immediate workflow inefficiencies but also sets the foundation for future advancements in data science. This paper explores the methodologies employed, key deliverables, and the broader implications for the discipline, ultimately demonstrating the project's potential to transform data-driven decision-making.

Background and Motivation

As organizations seek to harness insights from their growing data repositories, the demand for data-driven decision-making has surged across industries such as healthcare, finance, and retail. However, the rapid expansion of data science practices has revealed significant inefficiencies and challenges in workflow management, reproducibility, and ethical accountability. These challenges not only affect operational efficiency but also raise concerns about the reliability and fairness of AI-driven systems.

A key issue in data science is the lack of standardized practices and reusable solutions. Unlike software engineering, where pattern languages have been widely adopted to streamline development processes (Gamma et al., 1995), data science lacks a comparable framework. This absence could result in fragmented workflows, redundant efforts, and inconsistent outcomes. For example, a fundamental step in any data science project – data cleaning is frequently approached in an ad hoc manner, which leads to discrepancies that can affect the validity of subsequent analyses. Similarly, ethical concerns, such as algorithmic bias and lack of transparency, remain pervasive in AI applications, particularly in sensitive domains like hiring and healthcare (Mittelstadt et al., 2016).

The creation of a data science pattern language was conceived to address these gaps by drawing inspiration from established frameworks in software engineering and other disciplines where applying pattern language frameworks successfully and tailoring them to the unique challenges of data science. Our project developed a repository of reusable patterns that encapsulate best practices for common data science tasks such as data cleaning, feature engineering, and model evaluation using the CRISP-DM framework. By integrating these patterns with an AI-driven chatbot, data scientists apply them effectively within their workflows. To further evaluate the chatbot's effectiveness, we conducted a validation study comparing its classification of pattern applicability to that of a subject matter expert. This comparison provided insights into the chatbot's reliability as a decision-support tool and highlighted opportunities for future improvement in automated data science pattern application.

On an operational level, reusable patterns can significantly reduce redundancy, improve workflow efficiency and enhance reproducibility across projects. On an ethical level, incorporating principles of fairness, accountability, and transparency into the patterns makes data science practices with societal expectations and regulatory requirements. This dual focus positions the project as an important step toward addressing both the technical and ethical challenges facing the discipline.

Furthermore, the project fits into the increasing demand for tools that support real-time decision-making. By embedding the patterns within a chatbot, the project extends the usability of the pattern language framework and provides data scientists with on-demand guidance that reduces reliance on ad hoc expertise and fosters a more standardized approach. This innovation enhances individual productivity and promotes collaborative and consistent practices across teams.

LITERATURE REVIEW

Pattern Languages and Reusable Solutions

The concept of pattern languages originates from Christopher Alexander's work in architecture (Alexander 1979). This process has been successfully translated from the architectures of buildings to the architectures of computers (Coplien, 1996) and by Gamma et al. (1995) in *Design Patterns: Elements of Reusable Object-Oriented Software*. This book introduced the concept of reusable solutions for recurring problems, providing a structured approach to improving software development processes. The Alexandrian form used to document patterns in this project draws directly from Gamma's work, emphasizing clarity and standardization. Fortino (2005) extended this framework to technology management, and Eckstein (2000) extended it to education.

In data science, pattern languages remain an underexplored area, but their potential to address inefficiencies is significant. For example, Saltz et al. (2018) discuss the importance of codifying best practices in data science to improve project reproducibility and scalability. Their work underscored the need for structured frameworks, which influenced the decision to document patterns in a reusable, modular format.

Data Science Best Practices

Several key texts and papers provided insights into best practices in data science, forming the foundation for many of the patterns developed in this project. Donoho (2017) in *50 Years of Data Science* emphasized the importance of clean workflows and reproducibility in the data science lifecycle. This work showed the need for well-defined processes in areas like data cleaning, feature engineering, and model evaluation, directly influencing patterns such as *Data Cleaning Pipeline* and *Feature Engineering Framework*.

Another influential work was Kuhn and Johnson's *Applied Predictive Modeling* (2013), which provides a comprehensive guide to building and validating predictive models. Their emphasis on methods such as cross-validation, hyperparameter tuning, and the use of baseline models inspired several patterns, including *Cross-Validation Standard* and *Baseline Model Comparison*. Kuhn and Johnson's work also stressed the importance of understanding data distributions and the potential for concept drift, which became the foundation for the *Concept Drift Detection* pattern.

Ethical AI and Responsible Data Science

The integration of ethical principles into AI and data science workflows is a growing concern addressed by recent literature. Mittelstadt et al. (2016) in *The Ethics of Algorithms* provided a framework for understanding the societal implications of algorithmic decisions, particularly in addressing bias, transparency, and accountability. This work informed patterns like *Ethical AI and Bias Checking* and guided the chatbot's design to ensure it aligned with principles of fairness and interpretability.

Raji et al. (2020) in *Closing the AI Accountability Gap* discussed the importance of embedding fairness and bias mitigation into every stage of the AI lifecycle. This work reinforced the need to prioritize ethical considerations in the pattern library, particularly in domains like healthcare and finance, where the stakes are high.

Chatbot Development and Natural Language Processing

The development of the AI-driven chatbot drew heavily from literature on natural language processing (NLP) and conversational AI. Jurafsky and Martin's *Speech and Language Processing* (2020) provided a foundational understanding of NLP techniques, including intent recognition and response generation. This work guided the design of the chatbot's architecture and make sure it could interpret user queries and retrieve relevant patterns effectively.

Bickmore and Cassell (2005) in *Social Dialogue with Embodied Conversational Agents* explored the importance of user-centric design in chatbot development which demonstrated the need for an intuitive and interactive interface. These principles shaped the chatbot's conversational capabilities and ensure it provided actionable guidance in an accessible manner.

THEORETICAL DEVELOPMENT/MODEL

The project aimed to develop a pattern language specifically tailored for data science projects and their execution, leveraging the foundational work of past work by the authors on pattern languages and chatbot development, especially in using industry laws to create decision support systems (Fortino, 2024) and (Yu, 2024) The objective was to create a chatbot or a series of prompts that apply these patterns to data science, enabling more effective project execution and decision-making.

This proposed the discovery and validation of patterns using Generative AI, the aggregation of these patterns into a coherent language, and the development of tools to apply this language to well-known data science problems.

The Problem

The field of data science lacks a standardized, comprehensive framework for project execution and decision-making, similar to pattern languages used in other disciplines like architecture and software development. Current methods often fail to leverage the collective wisdom of industry laws and best practices in a systematic, easily applicable manner. There is a need for a structured, yet flexible, approach that can guide data scientists and project managers through the complexities of data science projects, ensuring that established patterns and industry laws are effectively applied to enhance project outcomes.

Theory

The theory behind this project is that a pattern language, as conceptualized by Christopher Alexander for architecture and adapted for various fields, can be effectively developed and applied to data science. By leveraging Generative AI tools to discover patterns from existing industry laws, validating these patterns with experts, and aggregating them into a coherent language, it's theorized that a comprehensive framework for data science project execution can be created. Furthermore, the theory posits that this pattern language can be operationalized through an AI-driven chatbot or series of prompts, enabling data scientists and project managers to effectively apply these patterns to specific data science problems, thereby improving project outcomes and decision-making processes.

Research Questions

1. Can a pattern language be effectively developed and applied to enhance data science project execution and decision-making?
2. Can an AI-driven chatbot or series of prompts based on a data science pattern language be developed to improve the efficiency and effectiveness of data science projects compared to traditional approaches?

Hypothesis

H1: A chatbot or series of AI-driven prompts based on a comprehensive data science pattern language can significantly improve the efficiency and effectiveness of data science project execution and decision-making, resulting in measurable enhancements in project outcomes, as demonstrated through simulations and expert evaluations.

METHODS

The creation of a data science pattern language employed a structured and iterative approach to develop a reusable pattern library and an AI-driven chatbot. The methodology was designed to address inefficiencies in data science workflows and provide actionable solutions. We detail the steps taken during the development process, including the design and validation of the deliverables and strategies to mitigate risks and challenges.

Patterns as Rules of Architecture

Christopher Alexander (1979) argues that architectural design rules may be codified through use of pattern languages. He calls the process the “the timeless way of building” (Alexander, 1977). He urges architects to build buildings that are alive: good to live in and to work in. This quality of aliveness or livability as apposed to “dead” or difficult to live in he says is hard to quantify, even to define. He calls it the quality without a name. But that quality can be recognized when it is present. Living architectures, that evolve gracefully, that serve people, that “feel” good have that quality.

Alexander demonstrates that the gateway to building with quality that lacks a name lies in the use of patterns. Patterns are essential elements of architecture. These patterns can be sought, discovered, documented, and utilized by other practitioners. One of Alexander's key contributions is teaching how to discover and document patterns (Alexander, 1979).

The process has been applied equally well to other areas of technology, especially ones that are organized in architectures. One could argue that data science is one such area, and data science as information and decision support systems architects can derive some benefit from using the Alexander techniques.

Development of the Pattern Library

Patterns are derived from the observation of industry forces in play, which eventually create problems. They, in turn, may find recurrent solutions. Industry laws are a fruitful source of well-known forces at play with known solutions. They are the wisdom of the elders of the industry and have accumulated over the years. To that end we used a set of data science industry laws identified and codified for another project (Huang, 2025).

The creation of the pattern library was a central focus of the project. The process involved the following steps:

Pattern Identification: Analyzed common challenges in data science workflows, such as data preprocessing, feature engineering and handling concept drift. Adapted established data science principles (e.g., Occam's Razor, Garbage In–Garbage Out) into actionable patterns. Using the 53 identifiable data science industry laws and the use of a Large Language Model, we identified 53 relevant patterns (See Appendix A)

Pattern Documentation: Each pattern was documented using the Alexandrine form:

Context: The specific scenario where the pattern is applicable.

Problem: The challenge the pattern addresses.

Forces: Conflicting factors that complicate the problem.

Solution: A detailed, step-by-step resolution.

Resulting Context: The outcome or changes after applying the pattern.

Examples: Real-world applications demonstrating the pattern's utility.

Consequences: Some results of applying the pattern

Related Patterns: Closely related patterns.

Pattern Validation: Presented the pattern library to domain experts for feedback on clarity, relevance, and applicability. Applied selected patterns to real-world data science problems, such as fraud detection and customer segmentation, to evaluate their utility.

Chatbot Design and Development

The AI-driven chatbot was designed to make the pattern library accessible and actionable for data scientists in real-time. The chatbot development process included:

Architecture Design: Adopted a modular design, separating components such as intent recognition, pattern retrieval, and response generation. This modularity facilitated easier troubleshooting and scalability. Selected Rasa, an open-source NLP framework for chatbot implementation due to its flexibility and support for customizable workflows.

Integration with Pattern Library: Developed a backend system to store and retrieve patterns dynamically based on user queries. Created mappings between common data science terms and patterns to enhance query accuracy.

Training and Testing: Trained the chatbot on a dataset of queries simulating real-world scenarios, such as “How do I handle missing data?” or “What is the best way to address bias in my model?” Conducted iterative testing to refine the chatbot’s ability to interpret user intent and provide relevant responses.

Validation and Testing: The project will involve rigorous testing of the developed chatbot or AI-driven prompts through simulations of real-world data science problems. Expert evaluations will be conducted to assess the tool’s effectiveness in applying the pattern language to various scenarios. The testing will involve comparing project outcomes and decision-making processes when using the tool versus traditional approaches.

To evaluate the alignment between human expert judgment and machine-driven classification, we conducted an **exploratory validation** comparing the **subject matter expert’s (SME) assessments** with the **LLM-generated classifications** of pattern applicability in the Carvana case study. The classification categories—**High, Medium, and Low applicability**—were assigned to each pattern by both the SME and the LLM. To measure the level of agreement, we used **Cohen’s Kappa coefficient**, a widely accepted statistical metric for inter-rater reliability, which accounts for agreement occurring by chance. Additionally, we applied **Weighted Cohen’s Kappa** to consider the ordinal nature of the categories, where discrepancies between “High” and “Medium” were treated as less severe than disagreements between “High” and “Low.” This analysis provided insight into how closely the chatbot’s classifications aligned with expert judgment and offered a preliminary assessment of the chatbot’s effectiveness in applying the Data Science Pattern Language to real-world problems.

Chatbot Testing: Tested the chatbot’s performance on a variety of queries to measure accuracy, usability, and user satisfaction. Collected feedback from testers to improve the chatbot’s response generation and error handling.

RESULTS

The creation of a data science pattern language yielded several impactful findings that present the utility and relevance of the deliverables in addressing inefficiencies in data science workflows. These findings are categorized into two main areas: the pattern library and the AI-driven chatbot, along with insights derived from their development and application. Each finding underscores the contributions of the project to standardizing practices and promoting ethical accountability in data science.

Pattern Library: Addressing Workflow Inefficiencies

The pattern library consists of some reusable best-practice solutions and emerged as a central output of the project. These patterns tackled common challenges in data science, such as data cleaning, feature engineering, model evaluation and bias mitigation. It provides structured solutions that are both practical and adaptable. We use the CRISP-DM (Shearer, 2000) framework of data science as a way to organize and apply the pattern to the appropriate phases of a data science project. Figure 1 shows the distribution of patterns top the framework.

CRISP-DM Phase	Patterns
Business Understanding	Cost-Benefit Tradeoff Analysis
Data Understanding	Collaborative Data Labeling, Data Debt Tracking, Data Encryption and Secure Access, Data Privacy Compliance Framework, Data Provenance Tracking, Data Validation Checks, Multi-Layer Data Security, Robust Outlier Detection
Data Preparation	Automated Feature Selection, Context-Specific Data Sampling, Data Augmentation Strategies, Data Cleaning Pipeline, Data Partitioning Framework, Dimensionality Reduction Techniques, Distributed Data Processing, Feature Engineering Framework, Hybrid Cloud Data Management, Missing Data Imputation, Resampling Techniques for Imbalanced Data, Synthetic Data Generation
Modeling	Active Learning Framework, Baseline Model Comparison, Collaborative Filtering Framework, Cross-Validation Standards, Ensemble Learning Techniques, Gradient-Based Optimization, Hyperparameter Optimization Loop, Incremental Learning Workflow, Modular Algorithm Design, Reward Engineering for Reinforcement Learning, Multi-Objective Optimization, Transfer Learning Framework
Evaluation	Anomaly Explanation, Automated Model Documentation, Ensemble Interpretability Workflow, Ethical AI and Bias Checking, Explainable AI in Production, Interactive Visualization Framework, Model Interpretability, Transferable Skills Documentation
Deployment	Automated Pipeline Testing, Concept Drift Detection, Continuous Integration for Data Science, Data Pipeline Automation, Deployment Environment Standardization, Dynamic Thresholding for Alerts, Feature Drift Management, Model Deployment Pipeline, Model Retraining Schedule, Performance Monitoring After Deployment, Real-Time Data Streaming Framework, Scalable Logging Framework

Figure 1 – The Discovered data science patterns classified into the CRISP-DM framework

One of the most significant findings was the effectiveness of using pattern languages to standardize workflows. By encapsulating best practices into a reusable format, the patterns streamlined tasks that are often performed inconsistently across projects. For example, the *Data Cleaning Pipeline* pattern provided a step-by-step guide for handling missing data, outlier detection and normalization which testers found to be highly replicable and efficient. Similarly, the *Feature Engineering Framework* pattern offered a structured approach to identifying and creating predictive features, reducing redundant effort.

The inclusion of ethical considerations within the patterns also proved impactful. Patterns like *Ethical AI and Bias Checking* and *Model Interpretability Standards* demonstrated the feasibility of integrating fairness and transparency into technical workflows. These patterns were particularly well-received by experts in regulated industries, such as healthcare and finance where ethical accountability is paramount. The feedback show that such patterns fill a critical gap in existing practices, aligning technical solutions with societal and regulatory expectations.

Chatbot: Real-Time Decision Support

The AI-driven chatbot was the second element of our efforts, designed to make the pattern library accessible and actionable for data scientists in real time. The chatbot demonstrated its value in providing on-demand guidance and promoting consistency in applying best practices. The chatbot's natural language processing capabilities enabled users to query the system in plain language, such as "How do I handle missing data?" or "What steps can I take to reduce bias in my model?" The chatbot retrieved and presented relevant patterns along with step-by-step instructions, making sure that users could immediately integrate the solutions into their workflows. During testing, users reported that the chatbot significantly reduced the time spent searching for solutions and improved their confidence in applying standardized practices.

However, the findings also revealed limitations in the chatbot's functionality. While it performed well with straightforward queries, it struggled with more complex or ambiguous questions such as those involving overlapping contexts (e.g., "How do I handle missing data in healthcare datasets while addressing bias?"). These challenges give the need for further refinement of the chatbot's NLP capabilities, particularly in handling multi-turn dialogues and nuanced queries.

User Feedback and Validation

Feedback from domain experts and testers provided valuable insights into the practical relevance of the two elements of our project. The patterns were praised for their clarity, comprehensiveness, and applicability across a range of use cases by several SMEs we consulted. Experts noted that the patterns not only reduced inefficiencies but also fostered reproducibility and collaboration by providing a common framework for approaching data science tasks.

The real-world application of selected patterns further validated their utility. For example, the *Concept Drift Detection Pattern* was applied to a predictive model in a retail context where it successfully identified shifts in data distributions and recommended timely adjustments. Similarly, the *Cross-Validation Standard* pattern improved model evaluation accuracy by ensuring consistent testing methodologies across projects.

The chatbot also received positive feedback for its ease of use and integration into existing workflows. Testers appreciated the conversational interface and the ability to access actionable guidance without interrupting their work. However, the feedback presents the importance of expanding the chatbot's training dataset to handle industry-specific queries and edge cases more effectively.

Results of a Validation Test

To observe whether the patterns were effective, reasonably complete, and applicable to real-world data science problems, we applied the framework to a case study. We have effectively used this process in the past as a source of expertly crafted problems and solutions for other projects (Yu, 2024). Academic case studies offered by Harvard and Ivey, to name a few of the more useful repositories, present carefully curated real-world scenarios along with the accompanying expert (the case writer) solutions.

Description of the Case

Carvana, an online used car platform, leverages data science to predict “kicks,” or vehicles with significant defects that are costly to the business. The case centers on their 2012 Kaggle competition aimed at developing predictive algorithms and explores challenges such as data quality, scalability, market adaptability, and balancing cost savings with customer experience.

CRISP-DM Phase	Recommendations
Business Understanding	<i>Transferable Skills Documentation</i> : Retain knowledge from system implementation for scalability and future improvements.
Data Understanding	<i>Interactive Visualization Framework</i> : Use interactive tools to explore dataset factors for predicting ‘kicks.’
Data Preparation	<i>Automated Feature Selection</i> : Streamline the process of identifying important features. <i>Dimensionality Reduction Techniques</i> : Simplify high-dimensional data for better performance. <i>Data Debt Tracking</i> : Log known data issues to reduce long-term costs.
Modeling	<i>Reward Engineering for Reinforcement Learning</i> : Optimize dynamic decisions around vehicle selection. <i>Collaborative Filtering Framework</i> : Explore creative approaches to predict vehicle viability using historical data.
Evaluation	<i>Ensemble Interpretability Workflow</i> : Build stakeholder trust by explaining ensemble model decisions. <i>Model Evaluation Workflow</i> : Structure model performance assessment effectively.
Deployment	<i>Automated Model Documentation</i> : Ensure transparency and reproducibility in model deployment. <i>Continuous Integration for Data Science</i> : Support seamless updates during model refinement. <i>Scalable Logging Framework</i> : Track operational issues efficiently in production environments.

Figure 2 – Recommended patterns for the Carvana case study.

Analysis of the Case

We applied our Data Science Pattern Language framework to analyze Carvana’s challenges. Through this process, we identified the patterns presented in Table 2, such as Dimensionality Reduction and Feature Engineering Framework for data preprocessing, and Model Deployment Pipeline for operational scalability. The resulting patterns address Carvana’s needs and provide structured solutions for their data science-driven operations.

Our process and the discovered patterns aligned effectively with the Carvana case. The patterns provided structured approaches to address key challenges such as data preprocessing, scalability, and ethical considerations. This application demonstrated the utility of our framework in offering practical, adaptable solutions for a dynamic and data-driven operational context.

Results of the SME vs. LLM Comparison

The analysis revealed a fair level of agreement between the SME and LLM classifications, with a Cohen's Kappa score of 0.186, indicating a level of alignment that exceeds chance but leaves room for improvement. The Weighted Cohen's Kappa score of 0.186 suggested that the chatbot's performance remained consistent even when considering the severity of disagreements. While this level of agreement does not imply perfect alignment, it validates the chatbot's ability to approximate expert judgment in pattern applicability assessments. However, as an exploratory validation using a single case study, this result should be interpreted as an initial indication rather than a definitive confirmation of the chatbot's accuracy. Future work should expand this analysis across multiple cases and industries to further assess and refine the model's predictive reliability. Despite these limitations, the findings suggest that LLM-driven classification can serve as a viable decision-support tool, assisting data science managers in triaging which patterns to prioritize for their teams.

DISCUSSION

Ethical Implications and Practical Contributions

One of the most significant findings was the impact of embedding ethical principles into the patterns. The emphasis on fairness, transparency, and accountability resonated strongly with testers, particularly in domains where decision-making can have profound societal consequences. For example, the *Bias Mitigation in Predictive Models* pattern was highlighted as a critical resource for teams developing AI systems in hiring and lending, where biased outcomes could lead to discrimination.

The project also demonstrated the feasibility of integrating ethical guidelines into real-time decision-support tools. By explaining its recommendations, the chatbot promoted transparency and trust among users, reinforcing the importance of ethical practices in data-driven workflows.

Scalability and Future Potential

The findings underscored the scalability of the deliverables. The modular design of the pattern library allows for the addition of domain-specific patterns while the chatbot's architecture supports integration with enterprise tools like Jira or Microsoft Teams. These features position the project as a foundation for future advancements, such as automating pattern discovery through machine learning or enhancing chatbot capabilities with advanced NLP techniques.

Implications For Data Science And Decision-Making

The creation of a data science pattern language has implications for data science and decision-making, particularly in enhancing workflow efficiency, promoting ethical practices, and facilitating informed decisions.

Standardization and Efficiency in Data Science Workflows

The development of a comprehensive pattern library addresses common challenges in data science, such as data cleaning, feature selection, and model evaluation. By providing standardized solutions, the library promotes consistency and efficiency across projects. It helps

data scientists to focus more on analysis and interpretation rather than reinventing solutions for recurring problems. This standardization is crucial for reproducibility and scalability in data science endeavors.

Integration of Ethical AI Practices

Incorporating ethical considerations into the pattern library helps data science practices fit with responsible AI principles and it addresses issues like bias mitigation and transparency. This integration is vital in regulated industries such as healthcare and finance, where ethical lapses can lead to significant consequences. By embedding ethical guidelines into standard workflows, the project fosters trust and accountability in AI applications.

Real-Time Decision Support through AI-Driven Chatbots

The implementation of an AI-driven chatbot provides real-time guidance in applying data science patterns and enhances decision-making processes. By leveraging natural language processing capabilities, the chatbot facilitates immediate access to best practices and help data scientists to make informed decisions swiftly. This real-time support is particularly beneficial in dynamic environments where timely decisions are critical.

Democratization of Data Science Knowledge

By making the pattern library and chatbot accessible, the project democratizes data science knowledge and allows practitioners at various skill levels to adopt best practices. This accessibility promotes a culture of continuous learning and improvement, it is essential for the advancement of the data science field.

Implications for Organizational Decision-Making

For organizations, the adoption of standardized patterns and real-time decision support tools can lead to more consistent and ethical decision-making processes. This alignment enhances operational efficiency and ensures that data-driven decisions adhere to ethical standards, thereby improving overall organizational performance.

FUTURE WORK

Expansion of the Pattern Library

Developing specialized patterns tailored to specific industries, such as healthcare, finance, and retail, can address unique challenges and regulatory requirements. Thereby increasing the library's applicability and effectiveness. Regularly updating the library to include patterns that leverage advancements in machine learning and artificial intelligence ensures that practitioners have access to state-of-the-art methodologies.

Enhancement of the AI-Driven Chatbot

Integrating more sophisticated natural language processing models can improve the chatbot's ability to understand and respond to complex, context-rich queries, thereby enhancing user interaction and support. Implementing support for multiple languages can broaden the chatbot's accessibility, catering to a diverse, global user base. Embedding the chatbot into popular data

science tools and platforms can provide seamless, real-time assistance within existing workflows, increasing its utility and adoption.

Automation of Pattern Discovery

Employing machine learning algorithms to analyze data science projects can facilitate the automatic identification and extraction of effective patterns, streamlining the expansion of the pattern library. Developing agentic systems that learn from user interactions and feedback can dynamically update and refine patterns.

CONCLUSIONS

The creation of a data science pattern language has made significant strides in addressing inefficiencies and ethical challenges within data science workflows. By developing a comprehensive pattern library and an AI-driven chatbot, the project offers standardized solutions and real-time guidance and promotes consistency and ethical compliance across various data science applications.

The pattern library serves as a repository of best practices and helps data scientists to tackle common challenges effectively. The AI-driven chatbot enhances this utility by providing immediate, context-aware support, facilitating the seamless integration of these patterns into daily workflows. This combination not only streamlines processes but also fosters a culture of continuous learning and adherence to ethical standards.

However, the project also encountered challenges, including resource constraints and the complexities of embedding ethical considerations into technical solutions. These experiences underscore the importance of multidisciplinary collaboration and iterative development in creating robust data science tools.

Looking forward, expanding the pattern library to include domain-specific solutions, it will enhance the chatbot's natural language processing capabilities and automating pattern discovery through machine learning are promising directions. Additionally, strengthening ethical AI practices and providing comprehensive user training will be crucial in ensuring the project's long-term success and relevance. The validation of the chatbot's pattern applicability classifications against a subject matter expert's assessments demonstrated a fair level of agreement, indicating that LLM-driven decision support tools can approximate expert judgment in guiding data science workflows. While this exploratory validation is limited to a single case study, it provides an initial indication that automated pattern classification can assist data science managers in prioritizing relevant methodologies, with further studies needed to refine and confirm these findings.

In conclusion, the creation of a data science pattern language represents a significant advancement in the pursuit of efficient and ethical data science practices. By offering practical tools that address both technical and ethical dimensions, the project contributes to the development of more responsible and effective data-driven decision-making processes.

REFERENCES

- Alexander, C. (1979), *The Timeless Way of Building*, Oxford University Press, Oxford.
Alexander, C. (1977), *A Pattern Language*, Oxford University Press, Oxford.

- Angluin, D. (1980). Finding Patterns Common to a Set of Strings. *Journal of Computer and System Sciences*, 21(1), 46–62.
- Barzen, J., & Leymann, F. (2014). Costume Languages as Pattern Languages. In *Proceedings of Pursuit of Pattern Languages for Societal Change – Preparatory Workshop 2014*.
- Coplien, J.O. (1996). *Software Patterns*, SIGS Books & Multimedia, New York.
- Eckstein, J., & Voelter, M. (2000). Learning to teach and Learning to learn. In *Proceedings of EuroPloP*.
- Falkenthal, M., Breitenbücher, U., Barzen, J., Fehling, C., & Leymann, F. (2014). Efficient Pattern Application: Validating the Concept of Solution Implementations in Different Domains. *International Journal on Advances in Software*, 7(3&4), 710–726.
- Fortino, A.G (2005), A Pattern Language for the Teaching and Practice of Technology Management, *International Journal of Technology and Innovation Management Education* 1(1): 2005, Senate Hall Academic Publishing.
- Fortino, A. and Zhang, W. (2024), AI-Enhanced Decision-Making: Integrating It Best Practices With Expert Insight, *Decision Sciences Institute Conference*, Phoenix, AZ Nov. 21-23, 2024.
- Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley.
- Grenander, U., & Miller, M. (2007). *Pattern Theory: From Representation to Inference*. Oxford University Press.
- Huang, S., Fortino, A., & Lowrance, R. (2025, March). DataSciLawBot: An LLM chatbot for the application of data science industry laws to business decisions. Paper presented at the Northeast Decision Sciences Institute Annual Meeting, Hershey, PA.
- Leymann, F. (2019). Towards a Pattern Language for Quantum Algorithms. In *Proceedings of the International Workshop on Quantum Technology and Optimization Problems* (pp. 218–230). Springer.
- Lichtendahl, K. C. Jr., & Holtz, P. (2019). *Carvana: IsBadBuy?* University of Virginia Darden School Foundation.
- Mumford, D. (2002). Pattern Theory: The Mathematics of Perception. *Proceedings of the International Congress of Mathematicians*, 1, 401–422.
- Nowak, A., & Leymann, F. (2013). Green Business Process Patterns – Part II. In *Proceedings of the 6th IEEE International Conference on Service Oriented Computing & Applications* (pp. 109–116). IEEE.
- Schuler, D. (2009). Choosing Success: Pattern Languages as Critical Enablers of Civic Intelligence. In *Proceedings of the Portland Urban Architecture Research Laboratory Symposium*.
- Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(1), 13–22.
- Wickham, H., & Grolemund, G. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media.
- Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., & Teal, T. (2017). Good Enough Practices in Scientific Computing. *PLOS Computational Biology*, 13(6), e1005510.
- Yu, Y., Fortino, A. and Felberbaum, A. (2024), InfoSecPilot: Navigating the Complex Landscape of Information Security with an AI-Powered Knowledge Management Chatbot, *Decision Sciences Institute Conference*, Phoenix, AZ Nov. 21-23, 2024.

APPENDIX A

Data Science Patterns Organized using the CRISP-DM Framework

U. Data Cleaning Pipeline

Problem: Raw datasets often contain errors, missing values or inconsistencies that hinder analysis.

Context : When preparing data for analysis or model building.

Forces : Balancing completeness and accuracy while managing time constraints.

Solution: Implement a standardized pipeline for cleaning, transforming, and validating data at the preprocessing stage.

Examples: Removing null values from customer transaction datasets. Standardizing date formats in time-series data.

Consequences: Improved data quality but may increase preprocessing time.

Related Patterns: Feature Engineering Framework, Data Validation Checks.

U. Feature Engineering Framework

Problem: Poorly designed features can lead to inaccurate models.

Context : During model building or enhancement phases.

Forces : Need for domain expertise versus the time constraints of generating meaningful features.

Solution: Use automated or semi-automated tools to extract and evaluate features systematically.

Examples: Creating interaction terms for logistic regression. Automating feature selection using SHAP values.

Consequences: Higher model accuracy but increased computational costs.

Related Patterns: Data Cleaning Pipeline, Model Interpretability.

U. Baseline Model Comparison

Problem: The lack of a benchmark makes it challenging to evaluate model improvements.

Context : When developing new models or refining existing ones.

Forces : Balancing simplicity with performance.

Solution: Always implement and evaluate a baseline model (e.g., mean predictor) before exploring complex algorithms.

Examples: Using a decision tree classifier as a baseline in binary classification problems. Comparing neural networks to linear regression for sales prediction.

Consequences: Better context for evaluation but adds initial development time.

Related Patterns: Model Evaluation Workflow, Cross-Validation Standards.

U. Cross-Validation Standards

Problem: Overfitting or underfitting due to poor evaluation strategies.

Context : During model training and testing phases.

Forces : Striking a balance between computational resources and robust evaluation.

Solution: Use k-fold cross-validation with appropriate stratification based on dataset characteristics.

Examples: Applying 10-fold cross-validation in imbalanced datasets. Stratified sampling for fraud detection models.

Consequences: Enhanced reliability but computationally expensive.

Related Patterns: Baseline Model Comparison, Model Evaluation Workflow.

U. Model Interpretability

Problem: Stakeholders often distrust black-box models due to a lack of transparency.

Context : When presenting or deploying machine learning solutions.

Forces : Balancing model complexity with interpretability.

Solution: Use explainability tools like LIME or SHAP to elucidate predictions.

Examples: Explaining feature importance in a random forest model. Visualizing decision boundaries in SVMs.

Consequences: Increased transparency but may complicate model deployment.

Related Patterns: Ethical AI and Bias Checking, Baseline Model Comparison.

U. Ethical AI and Bias Checking

Problem: AI systems can perpetuate or amplify biases present in training data.

Context : When training, testing, or deploying AI models.

Forces : Ensuring fairness while maintaining accuracy.

Solution: Implement tools to measure and mitigate bias, such as disparate impact analysis.

Examples: Mitigating gender bias in hiring algorithms. Analyzing fairness in credit risk models.

Consequences: Increased complexity but enhanced ethical compliance.

Related Patterns: Data Validation Checks, Model Interpretability.

U. Concept Drift Detection

Problem: Models degrade over time as data distributions change.

Context : Post-deployment monitoring of machine learning models.

Forces : Detecting and responding to drift without causing downtime.

Solution: Regularly monitor feature distributions and retrain models as needed.

Examples: Detecting concept drift in customer churn prediction models. Identifying seasonality changes in sales forecasting.

Consequences: Requires additional resources but ensures model reliability.

Related Patterns: Model Evaluation Workflow, Performance Monitoring After Deployment.

U. Data Validation Checks

Problem: Errors in input data can lead to inaccurate predictions or results.

Context : During data ingestion and preprocessing.

Forces : Balancing validation thoroughness with computational overhead.

Solution: Automate checks for anomalies, missing values, and outliers before processing.

Examples: Validating input ranges in sensor data. Detecting anomalies in financial transactions.

Consequences: Enhanced data reliability but additional preprocessing time.

Related Patterns: Data Cleaning Pipeline, Ethical AI and Bias Checking.

U. Performance Monitoring After Deployment

Problem: Models may fail silently, leading to poor decision-making.

Context : After deploying machine learning solutions.

Forces : Ensuring model accuracy while minimizing monitoring costs.

Solution: Establish performance dashboards that track metrics like precision, recall, and latency.

Examples: Monitoring click-through rates in recommendation systems. Tracking accuracy in sentiment analysis models.

Consequences: Better oversight but resource-intensive.

Related Patterns: Concept Drift Detection, Cross-Validation Standards.

10. Automated Feature Selection

Problem: Manually selecting features is time-consuming and prone to errors.

Context : During model optimization and testing.

Forces : Tradeoff between manual insight and automation efficiency.

Solution: Use algorithms like recursive feature elimination or mutual information ranking.

Examples: Employing RFE in predictive maintenance. Using mutual information for feature ranking in fraud detection.

Consequences: Faster feature engineering but potential loss of domain-specific insights.

Related Patterns: Feature Engineering Framework, Model Evaluation Workflow.

11. Incremental Learning Workflow

Problem: Re-training models from scratch is time-consuming and computationally expensive.

Context : When new data arrives or model updates are needed frequently.

Forces : Balancing resource usage with model accuracy.

Solution: Use incremental learning techniques to update models with new data without re-training entirely.

Examples: Updating recommendation systems with daily sales data. Adapting spam detection models as email patterns evolve.

Consequences: Reduced computational cost but potential for suboptimal global performance.

Related Patterns: Concept Drift Detection, Performance Monitoring After Deployment.

12. Hyperparameter Optimization Loop

Problem: Poor hyperparameter choices lead to suboptimal model performance.

Context : During the training phase of machine learning models.

Forces : Balancing the time required for optimization with achieving maximum performance.

Solution: Use techniques like grid search, random search or Bayesian optimization to automate hyperparameter tuning.

Examples: Optimizing the number of layers in a deep learning model. Tuning the regularization parameter in a logistic regression model.

Consequences: Improved model performance but increased training time.

Related Patterns: Cross-Validation Standards, Baseline Model Comparison.

13. Data Partitioning Framework

Problem: Poor data splits can result in misleading model evaluations.

Context : During data preparation for model training and testing.

Forces : Ensuring a representative split without leakage.

Solution: Use stratified splitting techniques for imbalanced datasets or k-fold cross-validation for robust evaluation.

Examples: Stratified sampling in fraud detection datasets. Splitting datasets by time in stock market prediction models.

Consequences: More robust models but potential preprocessing overhead.

Related Patterns: Cross-Validation Standards, Baseline Model Comparison.

14. Data Augmentation Strategies

Problem: Insufficient training data often leads to poor generalization.

Context : In scenarios with small or imbalanced datasets.

Forces : Balancing data variety with preserving original data integrity.

Solution: Use augmentation techniques such as oversampling, synthetic data generation, or transformations.

Examples: Generating synthetic images for deep learning. SMOTE oversampling in imbalanced classification problems.

Consequences: Increased dataset size but potential for unrealistic data.

Related Patterns: Feature Engineering Framework, Data Validation Checks.

15. Ensemble Learning Techniques

Problem: Single models often fail to capture complex patterns effectively.

Context : When improving predictive accuracy is critical.

Forces : Tradeoff between complexity and interpretability.

Solution: Combine multiple models (like bagging, boosting, or stacking) to improve performance.

Examples: Using random forests for classification tasks. Applying gradient boosting in regression models.

Consequences: Enhanced accuracy but increased computational costs.

Related Patterns: Model Interpretability, Performance Monitoring After Deployment.

16. Missing Data Imputation

Problem: Models struggle to perform with missing or incomplete data.

Context : During preprocessing and data cleaning phases.

Forces : Balancing imputation accuracy with maintaining data integrity.

Solution: Use imputation methods such as mean, median or predictive models to fill in missing values.

Examples: Imputing missing values in healthcare records. Using predictive models to estimate missing financial transaction data.

Consequences: Improved data usability but potential biases introduced.

Related Patterns: Data Cleaning Pipeline, Data Validation Checks.

17. Model Deployment Pipeline

Problem: Deploying models to production environments is error-prone and time-consuming.

Context : After model development and validation.

Forces : Balancing automation with deployment flexibility.

Solution: Implement CI/CD pipelines tailored for machine learning deployments.

Examples: Deploying a fraud detection model via Docker containers. Using MLOps frameworks to automate model versioning and deployment.

Consequences: Streamlined deployment but initial setup costs.

Related Patterns: Performance Monitoring After Deployment, Concept Drift Detection.

18. Dimensionality Reduction Techniques

Problem: High-dimensional datasets can cause computational inefficiency and overfitting.

Context : Before training machine learning models.

Forces : Balancing dimensionality reduction with retaining critical information.

Solution: Apply techniques like PCA, t-SNE, or autoencoders to reduce dimensionality while preserving key data characteristics.

Examples: Using PCA to compress feature space in image data. Applying t-SNE for visualizing high-dimensional text embeddings.

Consequences: Faster computation but potential information loss.

Related Patterns: Feature Engineering Framework, Baseline Model Comparison.

19. Automated Pipeline Testing

Problem: Errors in preprocessing or feature engineering can lead to invalid model outputs.

Context : During pipeline development and maintenance.

Forces : Balancing thorough testing with development speed.

Solution: Implement automated testing frameworks to validate each pipeline stage.

Examples: Writing unit tests for data cleaning scripts. Automating checks for feature consistency across datasets.

Consequences: Higher reliability but longer initial setup time.

Related Patterns: Data Cleaning Pipeline, Data Validation Checks.

20. Explainable AI in Production

Problem: Lack of explainability in AI systems can erode trust among stakeholders.

Context : During deployment and post-deployment monitoring.

Forces : Balancing model performance with transparency.

Solution: Integrate explainability tools (e.g. SHAP, LIME) to provide insights into model decisions.

Examples: Explaining credit score predictions to end-users. Visualizing feature importance in healthcare diagnostic tools.

Consequences: Increased stakeholder trust but potential performance tradeoffs.

Related Patterns: Model Interpretability, Ethical AI and Bias Checking.

21. Real-Time Data Streaming Framework

Problem: Batch processing delays can result in outdated insights for time-sensitive decisions.

Context : When handling real-time data feeds, such as IoT sensor data or live financial transactions.

Forces : Balancing low latency with reliable processing.

Solution: Implement stream-processing frameworks like Apache Kafka or Spark Streaming to process and analyze data in real time.

Examples: Monitoring network traffic for cybersecurity threats. Processing real-time stock market data for trading algorithms.

Consequences: Faster insights but requires significant infrastructure setup.

Related Patterns: Performance Monitoring After Deployment, Concept Drift Detection.

22. Robust Outlier Detection

Problem: Outliers can skew analyses or disrupt machine learning model performance.

Context : During data cleaning or preprocessing stages.

Forces : Balancing sensitivity in detecting anomalies with the risk of removing valid data points.

Solution: Use statistical methods (e.g. Z-scores) or machine learning techniques (e.g. isolation forests) to identify and handle outliers.

Examples: Detecting fraudulent transactions in financial datasets. Identifying faulty sensor readings in IoT systems.

Consequences: Improved data quality but potential removal of edge cases.

Related Patterns: Data Cleaning Pipeline, Data Validation Checks.

23. Distributed Data Processing

Problem: Large datasets exceed the memory or processing capacity of a single machine.

Context : When handling big data or computationally intensive workflows.

Forces : Balancing scalability with system complexity.

Solution: Use distributed computing frameworks like Apache Hadoop or Spark to parallelize data processing tasks.

Examples: Processing terabytes of e-commerce data for customer behavior analysis. Parallelizing genomic sequence analysis in bioinformatics.

Consequences: Faster processing at scale but increased setup and operational complexity.

Related Patterns: Real-Time Data Streaming Framework, Automated Pipeline Testing.

24. Transfer Learning Framework

Problem: Training deep learning models from scratch is resource-intensive.

Context : When tackling problems with limited data availability.

Forces : Balancing computational efficiency with model accuracy.

Solution: Use pre-trained models as a starting point and fine-tune them on the target dataset.

Examples: Fine-tuning BERT for text classification in a specific domain. Adapting ResNet for image recognition tasks in medical imaging.

Consequences: Faster training with better initial accuracy but requires expertise in model adaptation.

Related Patterns: Feature Engineering Framework, Dimensionality Reduction Techniques.

25. Data Pipeline Automation

Problem: Manually managing data pipelines increases the likelihood of errors and inefficiencies.

Context : When building scalable systems for regular data ingestion and preprocessing.

Forces : Ensuring reliability while reducing manual intervention.

Solution: Use tools like Apache Airflow or Prefect to automate pipeline orchestration.

Examples: Automating daily ETL processes for sales data. Scheduling regular feature extraction workflows in MLOps.

Consequences: Improved reliability and scalability but initial setup is resource-intensive.

Related Patterns: Automated Pipeline Testing, Distributed Data Processing.

U. Model Retraining Schedule

Problem: Models degrade in performance over time due to changing data distributions.

Context : In dynamic environments where data evolves frequently.

Forces : Balancing the need for up-to-date models with resource availability.

Solution: Establish a retraining schedule based on data drift thresholds or fixed intervals.

Examples: Monthly retraining of customer segmentation models. Weekly updates for real-time recommendation systems.

Consequences: Improved model performance but increases computational costs.

Related Patterns: Concept Drift Detection, Performance Monitoring After Deployment.

U. Synthetic Data Generation

Problem: Access to real-world data is often restricted due to privacy concerns or availability issues.

Context : When working with limited or sensitive datasets.

Forces : Balancing data realism with ethical and legal constraints.

Solution: Use generative models (e.g. GANs) or statistical simulations to create synthetic datasets.

Examples: Generating synthetic patient data for healthcare model training. Creating artificial customer profiles for marketing analysis.

Consequences: Expands training data but may introduce bias or unrealistic patterns.

Related Patterns: Data Augmentation Strategies, Ethical AI and Bias Checking.

U. Collaborative Filtering Framework

Problem: Personalization requires understanding user preferences in the absence of explicit feedback.

Context : When building recommendation systems or adaptive user interfaces.

Forces : Balancing computational complexity with prediction accuracy.

Solution: Use collaborative filtering techniques to infer user preferences from historical behavior.

Examples: Recommending movies on a streaming platform based on user ratings. Suggesting products in e-commerce based on purchase history.

Consequences: Enhanced personalization but can suffer from cold start problems.

Related Patterns: Ensemble Learning Techniques, Real-Time Data Streaming Framework.

29. Gradient-Based Optimization

Problem: Model training processes can converge to suboptimal solutions without efficient optimization techniques.

Context : During the training phase of machine learning or deep learning models.

Forces : Balancing convergence speed with avoiding local minima.

Solution: Use advanced optimization algorithms like Adam, RMSProp, or stochastic gradient descent.

Examples: Training neural networks with adaptive gradient techniques. Optimizing support vector machines using SGD.

Consequences: Improved convergence but can be computationally intensive.

Related Patterns: Hyperparameter Optimization Loop, Model Deployment Pipeline.

30. Data Provenance Tracking

Problem: Lack of traceability in data pipelines can lead to accountability issues.

Context : In highly regulated industries or collaborative environments.

Forces : Ensuring transparency while managing documentation overhead.

Solution: Implement tools or systems to track the origin, transformations, and lineage of datasets.

Examples: Using data provenance tracking tools in financial audits. Documenting pipeline stages in genomics research datasets.

Consequences: Enhanced traceability but additional overhead in maintaining logs.

Related Patterns: Data Validation Checks, Ethical AI and Bias Checking.

31. Data Encryption and Secure Access

Problem: Sensitive data requires protection from unauthorized access.

Context : When handling personal, financial, or medical data in compliance with privacy regulations.

Forces : Balancing security with usability in workflows.

Solution: Use encryption techniques and role-based access controls to safeguard data at rest and in transit.

Examples: Encrypting patient data in healthcare systems using AES. Implementing role-based access for financial transaction datasets.

Consequences: Enhanced security but added complexity in accessing data.

Related Patterns: Ethical AI and Bias Checking, Data Provenance Tracking.

32. Cost-Benefit Tradeoff Analysis

Problem: Projects often require resource allocation decisions with conflicting priorities.

Context : When evaluating the feasibility of new models or data pipelines.

Forces : Balancing cost constraints with expected benefits.

Solution: Use decision frameworks like ROI calculations or impact-effort matrices.

Examples: Evaluating the cost of GPU-based training for deep learning models. Assessing the tradeoff between automated pipelines and manual processes.

Consequences: Informed decision-making but requires robust data inputs.

Related Patterns: Performance Monitoring After Deployment, Automated Pipeline Testing.

33. Active Learning Framework

Problem: Labeling large datasets can be resource-intensive.

Context : When dealing with limited labeled data in supervised learning tasks.

Forces : Balancing labeling costs with improving model performance.

Solution: Use active learning to prioritize labeling the most informative samples.

Examples: Using uncertainty sampling in sentiment analysis datasets. Implementing active learning for image classification in medical diagnostics.

Consequences: Improved label efficiency but requires specialized techniques.

Related Patterns: Data Augmentation Strategies, Incremental Learning Workflow.

34. Transferable Skills Documentation

Problem: Teams often fail to document transferable skills gained during project execution.

Context : After completing major milestones or at project closure.

Forces : Balancing project speed with knowledge retention.

Solution: Maintain a skills repository for processes and technologies used in successful projects.

Examples: Documenting feature engineering techniques for fraud detection. Recording deployment strategies for cloud-based models.

Consequences: Improved onboarding and team efficiency but requires consistent updates.

Related Patterns: Data Provenance Tracking, Automated Pipeline Testing.

35. Multi-Objective Optimization

Problem: Projects often require optimizing multiple conflicting objectives.

Context : When tradeoffs between performance metrics (e.g. precision and recall) exist.

Forces : Balancing competing objectives while ensuring actionable insights.

Solution: Use optimization algorithms (e.g. Pareto optimization) to find balanced solutions.

Examples: Optimizing resource allocation in time-constrained projects. Balancing sensitivity and specificity in medical diagnostics.

Consequences: Holistic solutions but added computational complexity.

Related Patterns: Hyperparameter Optimization Loop, Gradient-Based Optimization.

36. Scalable Logging Framework

Problem: Inadequate logging makes debugging distributed systems challenging.

Context : During model training or production monitoring.

Forces : Balancing detailed logs with storage costs.

Solution: Use centralized logging frameworks like Elasticsearch, Logstash, Kibana or Fluentd.

Examples: Logging API requests for chatbot interactions. Monitoring training epochs in distributed deep learning systems.

Consequences: Improved debugging but increased storage needs.

Related Patterns: Automated Pipeline Testing, Distributed Data Processing.

37. Reward Engineering for Reinforcement Learning

Problem: Poorly defined rewards can lead to suboptimal reinforcement learning models.

Context : When designing RL systems for sequential decision-making tasks.

Forces : Balancing exploration with exploitation.

Solution: Iteratively refine reward functions to align with desired outcomes.

Examples: Designing rewards for self-driving car simulations. Refining rewards for optimizing warehouse robotics.

Consequences: Enhanced model alignment but requires iterative tuning.

Related Patterns: Gradient-Based Optimization, Incremental Learning Workflow.

38. Anomaly Explanation

Problem: Detected anomalies often lack clear explanations for stakeholders.

Context : In fraud detection, sensor monitoring, or other anomaly-related tasks.

Forces : Balancing interpretability with anomaly detection accuracy.

Solution: Use interpretable machine learning techniques to explain detected anomalies.

Examples: Explaining fraudulent transactions using SHAP values. Justifying anomaly alerts in industrial equipment monitoring.

Consequences: Improved trust in results but increased computational complexity.

Related Patterns: Model Interpretability, Robust Outlier Detection.

39. Resampling Techniques for Imbalanced Data

Problem: Imbalanced datasets can bias model training.

Context : During preprocessing for classification tasks.

Forces : Balancing minority class representation with dataset realism.

Solution: Use oversampling (e.g., SMOTE), undersampling or hybrid approaches to rebalance datasets.

Examples: Oversampling fraudulent transactions in financial datasets. Undersampling majority class in customer churn datasets.

Consequences: Improved model balance but potential overfitting.

Related Patterns: Data Augmentation Strategies, Feature Engineering Framework.

40. Feature Drift Management

Problem: Feature distributions may shift over time, reducing model reliability.

Context : Post-deployment monitoring of features.

Forces : Balancing retraining costs with model stability.

Solution: Monitor feature distributions and trigger retraining if drift exceeds predefined thresholds.

Examples: Detecting shifts in customer demographics in marketing models. Identifying changes in seasonal patterns for retail forecasting.

Consequences: Improved model consistency but requires robust monitoring systems.

Related Patterns: Concept Drift Detection, Performance Monitoring After Deployment.

41. Automated Model Documentation

Problem: Model documentation is often incomplete, hindering reproducibility.

Context : After finalizing model development or deployment.

Forces : Balancing thoroughness with automation.

Solution: Use tools like Mlflow or Sphinx to auto-generate model documentation.

Examples: Documenting feature importance and hyperparameters for regulatory compliance. Generating deployment notes for MLOps pipelines.

Consequences: Improved reproducibility but requires consistent tool usage.

Related Patterns: Data Provenance Tracking, Automated Pipeline Testing.

42. Interactive Visualization Framework

Problem: Static reports often fail to capture the complexity of data science outputs.

Context : During stakeholder presentations or exploratory data analysis.

Forces : Balancing detail with simplicity in visualizations.

Solution: Use interactive visualization tools like Plotly or Tableau to provide dynamic insights.
Examples: Building interactive dashboards for sales forecasting. Visualizing clustering results with user-defined filters.
Consequences: Better communication but requires advanced tools.
Related Patterns: Model Interpretability, Anomaly Explanation.

43. Multi-Layer Data Security

Problem: Simple security measures are often inadequate for sensitive data.
Context : In multi-user or distributed environments.
Forces : Balancing access flexibility with robust protection.
Solution: Implement a combination of encryption, access control, and audit logs.
Examples: Layered access for healthcare datasets. Multi-factor authentication for financial analytics systems.
Consequences: Enhanced security but increased complexity.
Related Patterns: Data Encryption and Secure Access, Ethical AI and Bias Checking.

44. Deployment Environment Standardization

Problem: Differences in environments lead to unpredictable model behavior.
Context : When transitioning from development to production.
Forces : Balancing development flexibility with production stability.
Solution: Use containerization tools like Docker to standardize environments.
Examples: Deploying fraud detection models in containerized environments. Standardizing feature pipelines for cloud-based systems.
Consequences: Improved reproducibility but requires familiarity with containerization.
Related Patterns: Model Deployment Pipeline, Automated Pipeline Testing.

45. Data Privacy Compliance Framework

Problem: Organizations risk non-compliance with privacy laws like GDPR and CCPA when handling sensitive data.
Context : When processing or sharing personal data.
Forces : Balancing data usability with legal and ethical requirements.
Solution: Implement data anonymization, encryption, and consent tracking to ensure compliance.
Examples: Anonymizing user identifiers in marketing datasets. Implementing consent-based data processing workflows.
Consequences: Ensures compliance but may reduce data granularity.
Related Patterns: Data Encryption and Secure Access, Ethical AI and Bias Checking.

46. Hybrid Cloud Data Management

Problem: Data scientists often need to work with distributed datasets across cloud and on-premise environments.
Context : When managing large-scale or sensitive datasets.
Forces : Balancing accessibility, performance, and security.
Solution: Use hybrid cloud frameworks to manage data across environments, ensuring accessibility while maintaining control over sensitive data.
Examples: Training models on public cloud resources while storing sensitive data on-premise. Using hybrid pipelines for regulatory compliance in cross-border data processing.
Consequences: Increased flexibility but higher infrastructure complexity.
Related Patterns: Distributed Data Processing, Multi-Layer Data Security.

47. Dynamic Thresholding for Alerts

Problem: Static thresholds for anomaly detection often lead to false positives or missed anomalies.
Context : When monitoring dynamic systems like real-time transactions or network traffic.
Forces : Balancing sensitivity with accuracy in dynamic environments.
Solution: Use adaptive algorithms to adjust thresholds based on real-time data characteristics.

Examples: Adjusting fraud detection thresholds during high-volume transaction periods. Dynamically setting error rate thresholds in server monitoring systems.

Consequences: Fewer false alerts but higher computational costs.

Related Patterns: Robust Outlier Detection, Real-Time Data Streaming Framework.

48. Ensemble Interpretability Workflow

Problem: Explaining ensemble model decisions is more challenging than single-model explanations.

Context : When using ensemble methods like random forests or gradient boosting.

Forces : Balancing model accuracy with interpretability for stakeholders.

Solution: Use ensemble-specific interpretability tools (e.g. SHAP for tree ensembles) to break down feature contributions.

Examples: Explaining credit approval decisions using SHAP values in gradient boosting. Visualizing feature importance across random forest models for customer churn predictions.

Consequences: Improved transparency but requires more computational resources.

Related Patterns: Model Interpretability, Anomaly Explanation.

49. Modular Algorithm Design

Problem: Algorithms designed for specific tasks are difficult to adapt to new requirements.

Context : When building reusable components for machine learning workflows.

Forces : Balancing generality with task-specific performance.

Solution: Design algorithms as modular components that can be customized or extended for new applications.

Examples: Creating a modular framework for different recommendation algorithms. Building a pipeline that supports multiple feature selection methods.

Consequences: Higher reusability but initial design complexity.

Related Patterns: Feature Engineering Framework, Automated Pipeline Testing.

50. Collaborative Data Labeling

Problem: Manual data labeling can be time-intensive and inconsistent across annotators.

Context : When labeling datasets for supervised learning tasks.

Forces : Balancing labeling quality with team productivity.

Solution: Use collaborative tools with built-in quality assurance (e.g., active consensus mechanisms) to streamline the labeling process.

Examples: Using tools like Label Studio for annotating medical images. Implementing majority-vote systems for resolving annotation disagreements.

Consequences: Improved annotation quality but requires coordination among annotators.

Related Patterns: Active Learning Framework, Synthetic Data Generation.

51. Continuous Integration for Data Science

Problem: Integrating new code or data into a collaborative data science project can introduce errors or conflicts.

Context: In multi-team environments with frequent updates to codebases or datasets.

Forces: Balancing rapid updates with system stability and reproducibility.

Solution: Implement continuous integration (CI) pipelines tailored to data science workflows, including automated testing of data transformations, model accuracy, and reproducibility checks.

Examples: Running automated tests for feature extraction scripts after each code commit. Verifying model outputs match expected metrics before integration.

Consequences: Faster integration and error detection but requires initial pipeline setup and maintenance.

Related Patterns: Automated Pipeline Testing, Model Deployment Pipeline.

52. Context-Specific Data Sampling

Problem: Data sampling methods often fail to capture domain-specific nuances, leading to suboptimal models.

Context: When selecting subsets of data for training or evaluation in diverse domains.

Forces: Balancing computational efficiency with data representativeness.

Solution: Use domain-aware sampling strategies that incorporate business logic, such as stratified sampling by customer segments or time-based sampling for seasonal effects.

Examples: Sampling transactions by region to account for geographic diversity in fraud detection. Stratifying survey data by demographic groups for marketing models.

Consequences: Enhanced model relevance but may require more domain expertise.

Related Patterns: Data Partitioning Framework, Data Validation Checks.

53. Data Debt Tracking

Problem: Accumulating unaddressed data quality issues increases long-term technical debt.

Context: In projects with evolving data pipelines or when datasets grow in complexity.

Forces: Balancing immediate deliverables with long-term maintainability.

Solution: Maintain a 'data debt ledger' to track known quality issues, their potential impacts, and plans for resolution.

Examples: Logging discrepancies in customer IDs across systems for future reconciliation. Documenting unaddressed missing values in historical datasets.

Consequences: Improved awareness and planning but requires ongoing updates and prioritization.

Related Patterns: Data Provenance Tracking, Data Validation Checks.

Development and Structure of an Introductory Graduate Course on AI for Business

AJ Raven
Department of Decision Sciences and MIS
LeBow College of Business
Drexel University
ar3783@drexel.edu

Abstract

This paper discusses the development and structure of an introductory course in Artificial Intelligence (AI) for business students. The course covers not just the business aspects of AI but also the underlying technologies. It includes both theory and applied work and is intended to make graduate students into leaders of AI development in their organizations. AI is considered in the broad sense. Not just GenAI, but also machine learning and robotic process automation. The paper covers how the course is being developed and will be kept up to date, and what the structure of the course will be. This is not a paper about how to use AI in teaching. It is instead about how to teach AI. The topic is AI in organizational contexts.

Course Structure

We need to teach many of our students about the basics of artificial intelligence (Shahid and Mishra, 2024). A student who has run a neural network through an end-to-end project will have a much better understanding of the possibilities and limitations of AI than a student who has only used ChatGPT to create a report outline, correct their grammar and generate the occasional picture.

We already teach students how to design well-structured relational databases using ERDs and normalization. But in reality only a small fraction of our students will ever work directly with the structure of a database. Yet it is valuable to the students to learn those skills because they have to be able to work with the people that do that work, and they need to understand the importance of for instance data cleaning.

We are already starting to see specialized AI courses, e.g. for deep learning, for decision-making. The course proposed here is an introductory graduate course that will afterwards be adapted as an undergraduate course.

From what is being published in the media (The Economist, 2024; Bousquette, 2024) it is clear that at this point only the companies that create neural networks or provide the infrastructure to run them are making money off AI. Regular businesses that are trying to cut cost or create value are struggling to get back any returns on their AI investments. Yes, code can be written much faster and marketing plans appear to be a popular target for successfully using GenAI, but it is essential for students graduating from our business schools to be able to see how AI can impact the bottom line.

AI comes with a lot of new terms. By learning the meaning of these terms students develop a vocabulary that will allow them to stay up to date with new developments.

The course assumes only a very basic understanding of information systems and AI. The ability to code in any language is not required.

High level learning goals

Be able to understand enough about AI to:

- *Make informed investment decisions about AI systems development.*
An organization that wants to for instance create its own Large Language Model will have to pay for the basic foundation model, carefully select which parts of its own corporate knowledge it wants to add to the LLM and pay for running the LLM. Students should be able to compare and evaluate different AI technologies for a specific business scenario and then do a cost-benefit analysis of deploying an AI system.
- *Understand the difference between using AI for efficiency (cost reduction) and for value/revenue creation (new products)*
Many companies are experimenting with GenAI. Some are seeing some reduction in costs in certain positions, a common example being software developers using GenAI to generate parts of their code. However, companies find it very hard to improve their products or services in such a way that they add value for customers. This in turn means that it is very difficult to show a return in the investment in AI for products or services.
- *Guide and track development of AI systems and infrastructure.*
The leadership of organizations has to keep up with all the new AI developments and be able to see which of these developments could be of importance for the organization, versus which are simply noise.
- *Continue to develop understanding of AI.*
Once the main concepts of AI (see Appendix A) are understood and placed into a theoretical framework it becomes much easier to learn about new developments and their impacts.

Technical Topics

Theory: Types of neural networks, machine learning, Robotic Process Automation (RPA)/intelligent agents, what they are used for.

Theory: Data streams for machine learning

Hands on with machine learning. This and creating neural networks requires the teaching of basic coding in the Python language.

Hands-on with a convolutional neural network

Hands-on with a recurring neural network (this can be training a very simple LLM, or building on a small foundation model).

Hands-on with prompt writing for GenAI

Business Topics:

Ethical, legal and environmental risks of AI; regulation

Return on AI investments

Use cases: software development, law, accounting, etc.

Selection of data for LLMs

Choices of LLMs (e.g. small vs. mid vs. large)

How the course is being developed and how it will be kept up to date

Developments in decision making tools like linear programming and decision trees are slow, but the information technology tools we use have always been changing. A typical MIS textbook is outdated before it is released, recently causing some publishers to switch to ‘Evergreen’ on-line textbooks that are updated more frequently but are still outdated on the latest technologies. Enter artificial intelligence and the rate of change in new tools and their applications is breathtaking.

One of the very challenging aspects of teaching AI is that there is so much to learn and keep up with. There are a number of steps that an instructor can take to keep AI knowledge fresh.

How AI instructors can stay up to date on new developments

- Search for publications on AI. In publications like Gartner, Harvard Business Review, CIO, the Economist and daily newspapers there is a plethora of articles on AI.
- Try to run neural networks that have been created by others. Many developers discuss their work through YouTube end-to-end projects. Their code is published in places like GitHub.
- Learn about AI theory. Even though many aspects of AI are constantly evolving, the theories behind it don’t necessarily change. Learning can happen through books, GenAI and on-line courses.
- Speak with colleagues, industry contacts, students.
- Review syllabi for AI courses from other departments or universities
- Track the reliability of GenAI output. Because the results from GenAI are based on probability it is not yet very reliable. Despite the improvements of scientific problem solving by GenAI, it is generally still not very good. One way to test it is to give it a Linear Programming Problem. Another is to have it do a text analysis, asking it for instance to do counts of how often someone speaks across multiple transcripts of conversations.
- Be an exemplar of the usefulness of AI by using it in research, service and teaching. Additionally, try out new consumer-level tools and demo them in the classroom. The demos also provide a platform for students to discuss how they use AI.
- Follow use cases of GenAI and Robotic Process Automation (RPA). As the reliability of GenAI improves, more and more companies are going to use Chatbots as an interface between people and transactions, or even for complete automation. For instance, a customer can interact with a chatbot to change their phone plan or to schedule a return for something that was ordered on-line. Early use cases have been ordering at fast food drive through restaurants. The combination of GenAI and RPA is also known as intelligent agents. On the technology side, tools like Microsoft’s Azure Logic Apps and AWS Step Functions allow us to connect OpenAI LLM components with more traditional tools, for instance through APIs.

How the university is supporting development of the course

At the author's university the college of business each year appoints several faculty fellows for a two-year period. The college provides them with structured support as giving them a course release so that they can focus on curricular innovation. The current focus is on GenAI.

An issue holding back innovation is that employees and students are only allowed to use approved technologies. For generative AI, the only approved system is Microsoft Co-pilot. It is provided free of charge to all employees and students. The restrictions are in place because of concerns about data privacy and cost. The latest AI tools tend to be expensive and raise a lot of concerns about protection of the data that is entered into the LLMs (Litan, Elliot, Sicular, Vashisth & Rigon, 2024). Also, tools that are currently relevant may be renamed, changed, or removed within six months. At the same time new tools become available on a continuous basis. A procedure is in place to request approval for new tools but it is time consuming.

References

Bousquette, I. (2024) It's Time for AI to Start Making Money for Businesses. Can It? Wall Street Journal.

Brethenoux, E. and B. Yan (2024) Defining AI and Setting Realistic Expectations. Gartner Inc.

Litan, A., Elliot, B., Sicular, S., Vashisth, S. & Rigon, G. (2024). Applying AI – Governance and Risk Management. Gartner, Inc.

Shahid, A. R., & Mishra, S. (2024). A Framework for a Master's in Applied Artificial Intelligence Program in Computer and Information Systems Discipline. Journal of Information Systems Education, 35(4), 495-511.

The Economist (2024). Artificial Intelligence is Losing Hype.

Appendix A: Main Concepts, Building Tools and Consumer Products

General Concepts

GenAI

Machine Learning

Reinforcement Learning

Deep Learning

Hallucination

Unsupervised Learning

Supervised Learning

Parameter

Prompt Writing/Engineering

Robotic Process Automation

One Hot Encoding

GPU/TPU/NPU/CPU

Convolutional neural networks (CNN)

LLM (Large Language Models)

Recurrent neural networks (RNN)
Prompt Engineering
Generative Adversarial Network (GAN)
Tokenization
Guard Rails
Red Teaming

AI Ethics

Explainable AI
General Artificial Intelligence (AGI)
Machine Learning Bias
Cultural Impacts
Toxicity
Zero Data Retention

AI Building Tools

Python
Tensorflow
PyTorch
Keras
Numpy

Main Companies and their Consumer Products

OpenAI: ChatGPT
Google: Gemini
Anthropic: Claude
Microsoft: CoPilot
Perplexity AI: Perplexity
Meta: Llama
Mistral: Codestral
DeepSeek: DeepSeek R-1

Tackling the Ambiguity Challenge with Generative Artificial Intelligence: AICMA, A Framework for Identification, Classification and Mitigation of Ambiguity

Kushal Patel, Rutgers University, kp1264@scarletmail.rutgers.edu

Jim Samuel, Rutgers University, jim.samuel@rutgers.edu

Alexander Pelaez, Hofstra University, alexander.pelaez@hofstra.edu

Shreyash Kalal, Rutgers University, ssk241@scarletmail.rutgers.edu

Abstract

In this era of artificial intelligence (AI), ambiguity presents a significant challenge for information and communication management, often leading to misinterpretations and inefficiencies. The rise of generative AI (Gen AI) has further amplified this issue by producing text that is often unclear or open to multiple interpretations, which could impact decision-making in many critical areas. To address these challenges, we propose **AICMA**: a framework for **AI-driven Identification, Classification, and Mitigation of Ambiguity**. The AICMA framework consists of three core stages: **Identification**, which determines whether a given sentence is ambiguous or not; **Classification**, which classifies a sentence based on ambiguity level (High – Ambiguous, Low – Ambiguous, or Not – Ambiguous); and **Mitigation**, which utilizes large language models (LLMs) to adapt and regenerate ambiguous sentences for enhanced clarity while preserving their original intent. By improving textual interpretability, AICMA offers significant value across various domains, including education, healthcare, policy-making, and beyond. This framework is based on our theoretical framing of ambiguity, particularly in AI-generated text. It has the potential to contribute to developing more robust and reliable AI systems that will produce clearer and more interpretable outputs. Its adaptability will allow it to be integrated into existing AI systems, making it a versatile framework for developers and researchers aiming to enhance information and communication effectiveness. Ultimately, AICMA represents a significant step forward in addressing the complexities of ambiguity in AI-generated text, paving the way for more transparent and effective AI-driven text and speech generation solutions, from conversational agents to critical decision-support systems.

Keywords: Language Ambiguity, AICMA framework, Ambiguity identification, Ambiguity classification, Ambiguity mitigation, Generative AI, Large Language Models, Llama-3-8B, Natural Language Processing

1 Introduction

Ambiguity pervades nearly all facets of human experience, manifesting as uncertainty, vagueness, or interpretational inconsistency across communication, decision-making, and analysis. It is characterized as uncertainty about the probabilities with which outcome can occur. Language can lead to many ambiguous utterances. A word or a phrase may, in its relationship to other words or phrases, force the reader or listener to attempt to construe its meaning in multiple ways. It is possible for a sentence on its own to be vague, but it could be disambiguated within a certain frame of reference (Kadłub, 2017). We see that in our day-to-day communication, lexicon-semantic concepts often cause ambiguity, namely homograph, homonym, homophone, and polysemy (Ovu, 2016).

The existence of ambiguity presents a challenge to decision-makers as it eliminates the ability to apply standard optimization approaches, such as those based on calculating the objective expected values of alternative actions. Ambiguity arises in most strategically important decisions in some form because of the genuine limits on the decision-makers' rationality and the information available about the alternatives and the future (Arend, 2020). Investment decisions are decisions under uncertainty that are subject to both risk and ambiguity (Knight, 1921).

The necessity of tackling ambiguity arises from its propensity to destabilize processes that rely on precision and clarity. Many processes rely on precision and clarity, most of which could be disrupted due to uncertainties in probabilities. For example, misleading and ambiguous policy statements or market data analysis could lead to specious investment decisions. Equivocal policy language may cause polarization, undermine governance, and fuel misinformation (Bloom et al., 2007). In politics, ambiguous policy language may cause polarization, undermine governance, and fuel misinformation (Sunstein, 2005). Healthcare decisions based on ambiguous diagnoses or instructions can jeopardize patient outcomes, while in new media, ambiguous narratives amplify the spread of misinformation (Vosoughi et al., 2018). Similarly, ambiguity in interpersonal or business communications frequently leads to misaligned expectations and strained relationships, undermining collaboration and efficiency.

Research shows that, when given a choice between two options differing in their degree of ambiguity, humans tend to prefer the less ambiguous option. Humans often struggle with ambiguity due to cognitive biases, limited processing capacity, and the inherent subjectivity of interpretations. Research in psychology suggests that individuals exposed to high levels of ambiguity exhibit increased stress and decision fatigue, reducing their ability to make sound judgments (Curley et al., 1986).

The ambiguity challenge becomes even more complex in today's context, where generative artificial intelligence (AI) systems introduce unprecedented opportunities and risks. While generative AI excels in content creation, it often amplifies ambiguity by producing outputs susceptible to hallucinations, misaligned contexts, and varying interpretative meanings (Brown et al., 2020; Bommasani et al., 2021). These issues necessitate a robust framework to systematically identify, classify, and mitigate ambiguity, ensuring that AI-driven systems enhance clarity rather than exacerbate confusion. This paper introduces the "AICMA" framework, which leverages generative AI to address ambiguity challenges, filling critical gaps in the existing landscape of ambiguity management.

2 Literature Review

Knowledge ambiguity has been identified as a major obstacle to effective knowledge transfer and the implementation of overall knowledge management systems. Knowledge ambiguity comprises

two types of ambiguity: component ambiguity and causal ambiguity. Component ambiguity is uncertainty about knowledge content, whereas causal ambiguity is uncertainty about how to use the knowledge (Law, 2014). Causal ambiguity relates to ambiguity as to how organizational actions and results, inputs and outcomes, or competencies and advantages are linked. Causal ambiguity is important because of its implications for organizational performance (Konlechner & Ambrosini, 2019).

Investment decisions are decisions under uncertainty that are subject to both risk and ambiguity. Events for which the future outcome is unknown but the underlying distribution is known are referred to as “risky” (Knight, 1921). The Knightian uncertainty, or ambiguity as it is called by (Ellsberg 1961) and (Camerer and Weber 1992), is distinct from risk and describes events for which not only the future outcome but also the underlying distribution is unknown (Kostopoulos et al., 2021). Increased risk—the uncertainty over known possible outcomes—leads firms to decrease leverage. Conversely, greater ambiguity—the uncertainty over the probabilities associated with the outcomes—leads firms to increase leverage. Using a theoretically motivated measure of ambiguity, the empirical analysis presents evidence consistent with the prediction that ambiguity is positively related to leverage and shows that ambiguity has an important and distinct impact on capital structure choice (Izhakian et al., 2017).

Sentiment analysis has been widely used to weave through information overload and ambiguity to understand human behavior and tendencies in numerous domains, including finance and healthcare (Garvey et al., 2021; Samuel et al., 2020; Rahman et al., 2021; Ali et al., 2021; Samuel et al., 2020a, 2020b). Ambiguity weakens the predictive ability of sentiment on stock returns, whereas standard uncertainty increases it. Lower market participation during periods of high ambiguity is the likely driver of this effect (Wagner & Wei, 2024). Greater ambiguity will lead firms to decrease equity, increase debt, and enter financial distress earlier. Cash holdings are retained longer if the investors’ ambiguity aversion bias is sufficiently large, while cash holdings become less attractive when the combined impact of ambiguity and ambiguity aversion is relatively low (Agliardi et al., 2016).

Normalizing mentions of medical concepts to standardized vocabularies is a fundamental component of clinical text analysis. Ambiguity—words or phrases that may refer to different concepts—has been extensively researched as part of information extraction from biomedical literature. Distinction of ambiguity helps identify directions for advancing medical concept normalization research (Newman-Griffis et al., 2020). Ambiguity and vagueness in clinical practice guidelines reduce the likelihood of clinician adherence. They lead to inconsistent interpretation and, in turn, to inappropriate practice variation and medical errors. Resolving ambiguity and vagueness is an essential step in the computerized implementation of clinical practice guidelines. Successful resolution of ambiguity and vagueness requires an understanding of their characteristics, yet ambiguity and vagueness have not been differentiated, classified, and described in the medical context (Shlomi Codish & Shiffman, 2024).

In most medical decisions, probabilities are ambiguous and not objectively known. Empirical evidence suggests that people’s preferences are affected by ambiguity. Health economic analyses generally ignore ambiguous preferences and assume that they are the same as preferences under risk. For health gains, ambiguity preferences and risk preferences were indeed the same. For health losses, they differed, with subjects being more pessimistic in decisions under ambiguity. Utility and loss aversion were the same for risk and ambiguity. Implying that reducing the clinical ambiguity of health losses has more impact than reducing the ambiguity of health gains, that utilities elicited with known probabilities may not carry over to an ambiguous setting, and that ambiguity aversion may impact the value of information analyses if losses are involved (Attema et al., 2018).

In an acute crisis, it can be difficult to know exactly what has happened and how to respond. Sometimes, it's not a lack of information that is the biggest problem, but that huge amounts of information on social media create ambiguity. During such times, it is particularly important to be critical of sources, as both the behavior of users and the structure of the platforms contribute to this lack of clarity (Eriksson Krutrök & Lindgren, 2022). A news headline is an integral part of the journalistic article. Some journalists, intentionally or unintentionally, employ ambiguity when constructing news headlines. Ambiguity could positively affect the readers and motivate them to continue reading the news story. Conversely, it could also lead to misunderstanding and frustration (Khalifa, 2018). Ambiguity for advertising creativity and the effectiveness of persuasion can be termed strategic ambiguity. Present advertisers often deliberately design ads to spawn ambiguity, especially for certain product categories such as fashion and alcoholic drinks (Han & Choi, 2015).

Politicians generally have two motives: they wish to hold office as long as possible and wish to implement their preferred policies. Thus, they face a trade-off between the policies that maximize their chances of reelection and their most preferred policies (or the policies most preferred by the constituency that they represent) (Alesina & Cukierman, 1987). Ambiguity is often a double-edged sword that provides an opportunity to further a democratic agenda in the face of institutional barriers, but with that comes a more complicated and unclear causal pathway that connects citizens to decisions that impact public goods and services (Fowler, 2023). Using voter uncertainty to an advantage. Ambiguity is equated with voter uncertainty: a politician takes an ambiguous stance on an issue by announcing to the voters a lottery over the positions he might take on that issue (Dacey, 1979).

Research on ambiguity can be traced back to classical philosophy, where it was examined in relation to logical reasoning and truth. Aristotle's "Categories" explored the role of linguistic ambiguity in philosophical arguments (Thorp, 2016). In the 20th century, Wittgenstein's "Philosophical Investigations" delved into how language games create and resolve ambiguity (Wittgenstein, 1958). Traditional methods to address ambiguity include formal logic, probabilistic reasoning, and context-based disambiguation. Semantic analysis and syntactic parsing techniques in linguistics have been employed to resolve ambiguities in textual communication (Chomsky, 1957). Decision theory has leveraged Bayesian models to address ambiguities in uncertain environments (Savage, 1954). Frameworks like contextual bandits and reinforcement learning aim to resolve ambiguity in dynamic systems by iteratively refining decisions based on feedback (Sutton & Barto, 2018).

Contemporary solutions often combine human expertise with algorithmic tools. For example, legal and financial sectors utilize expert systems that incorporate predefined rules to disambiguate complex documents. In NLP, pre-trained models such as BERT and GPT attempt to resolve textual ambiguities through contextual embeddings (Devlin et al., 2019). However, these approaches are not immune to errors, especially in edge cases where nuanced human judgment is required. Recent advancements in AI have introduced novel methods for tackling ambiguity. Generative AI models are used to create paraphrases, clarifying ambiguous statements by presenting multiple interpretations. Hybrid approaches, combining symbolic reasoning with machine learning, aim to improve the reliability of AI-driven systems in addressing ambiguity (Lake et al., 2017). Despite these developments, challenges such as hallucination, context misalignment, and lack of accountability persist, limiting their efficacy.

Efforts to reconcile ambiguity and clarity focus on interpretability and transparency in AI systems. Techniques such as attention visualization in Transformer-based models provide insights into how these models resolve ambiguity (Vaswani et al., 2017). Additionally, human-AI collaboration frameworks emphasize using AI as a support tool rather than a replacement for human judgment,

thereby mitigating risks associated with ambiguity amplification. These approaches underscore the need for holistic frameworks that integrate technological precision with human contextual understanding.

The literature highlights a significant gap in comprehensive frameworks designed to address ambiguity across domains, particularly in the context of generative AI. By building on existing methods and leveraging the unique capabilities of generative AI, the proposed “AICMA” framework aims to establish a systematic approach to identifying, classifying, and mitigating ambiguity, thereby advancing theoretical understanding and practical applications in ambiguity management.

3 Data

Our research team created a dataset to study ambiguity. The dataset and examples used in this study are a manually selected subset of our main ambiguity dataset. The criteria for the selection of the subset was manually determined in alignment with our conceptual narrative in this introductory paper. The main full dataset was generated using large language models (LLMs) and includes sentences representing various levels and types of ambiguity, such as lexical, syntactic, referential, and scope ambiguity. The dataset includes sentences with corresponding labels: *High – Ambiguous*, *Low – Ambiguous*, and *Not – Ambiguous*. We have labeled the data based on classification by human experts who have significant experience in NLP, computational linguistics, and LLM validation. Hence the labeling is fairly reliable. Although different people can interpret the same sentence in various ways, our classification is not about the meaning but about the presence or absence of clarity about the meaning. The dataset for this research was specifically tailored to ensure relevance to a critical real-world context where ambiguity can have significant consequences. It is not our intention in this paper to present a comprehensive analysis of the main ambiguity dataset, which is being used for our future research.

4 Methodology

The research focuses on using **AI to identify, classify, and mitigate ambiguity (AICMA)** in text. Here, we introduce the AICMA framework, which consists of three core stages:

1. **Identification:** Determines whether a given sentence is ambiguous or not.
2. **Classification:** Classifies a specific ambiguity level (*High – Ambiguous*, *Low – Ambiguous*, or *Not – Ambiguous*).
3. **Mitigation:** Utilizes generative AI to rephrase by regenerating ambiguous sentences for enhanced clarity.

The framework as a whole employs an iterative loop to manage ambiguity by systematically processing information using AI for identification, classification, and mitigation of ambiguity (Figure 1). The framework is particularly tailored to address the critical need for precision and clarity in communication.

This flowchart, as shown in Figure 1, outlines a systematic process for identifying, classifying, and mitigating ambiguity in text using a language model (LLM). The process begins with uploading the input data. The model performs two key tasks: ambiguity identification and ambiguity classification. In the identification phase, the LLM identifies whether the text is ambiguous or not. Simultaneously, the classification phase categorizes the level of ambiguity, such as *High – Ambiguous*, *Low – Ambiguous*, or *Not – Ambiguous*. Next, a decision point evaluates the results

from both stages. If the text is identified and classified as “Not – Ambiguous,” the system directly displays the original sentence to the user. However, if

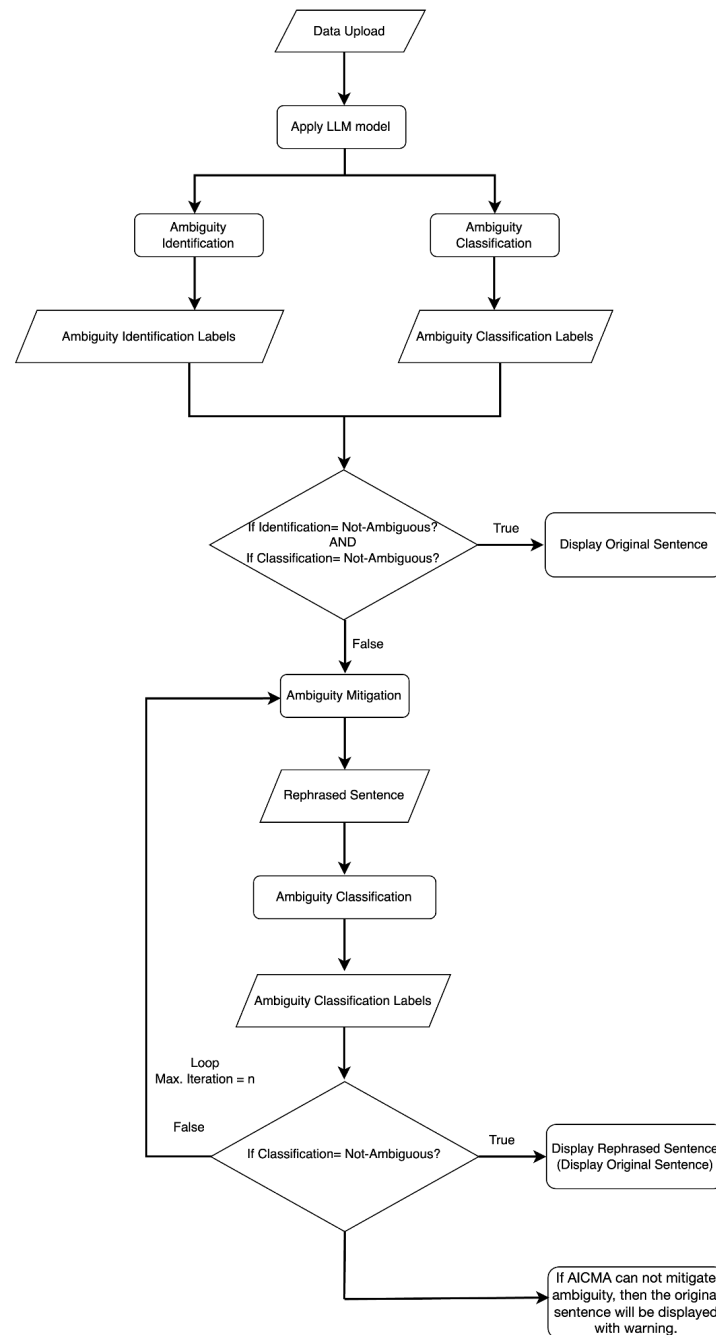


Figure 1: Step-by-step flowchart of the AICMA framework

ambiguity is detected in either the identification or classification stage, the process moves to an ambiguity mitigation step, where methods are applied to resolve the ambiguity. This involves regenerating the ambiguous sentence to enhance clarity. This mitigation process continues until

the regenerated sentence is classified as “Not – Ambiguous” or the loop is run for a maximum number of iterations in case the model fails to mitigate the ambiguity. Finally, if the ambiguity is mitigated, the AICMA displays both the original and regenerated sentences. The original sentence is displayed in a smaller font to maintain transparency and user awareness, allowing the user to make an informed judgment. This approach ensures that ambiguous text is systematically addressed, leading to clearer and more effective communication. If the AICMA fails to mitigate the ambiguity, it displays the original sentence with a warning.

4.1 AICMA – Identification of Ambiguity

Prompt engineering techniques were employed to enable the LLM to assess ambiguity in sentences. We iterated manually and experimented with prompts to arrive at a feasible point in accuracy for our dataset. Our experimentation revealed that prompts will need to be engineered based on which LLM is employed and the domain from which the data is being drawn. Table 1 shows a set of examples with Llama-3-8B (AI@Meta, 2024) and the variance between the human-annotated “Actual Label” (Gold standard) and the AI-assigned label “Identified Label.”

Table 1: Sentences with their actual and identified labels using Llama-3-8B

Sr No.	Sentence	Actual Label	Identified Label
1	The chicken is ready to eat.	Ambiguous	Ambiguous
2	I saw the man with the telescope.	Ambiguous	Ambiguous
3	The waiter served the meal with a smile.	Ambiguous	Not – Ambiguous
4	She found the address on the back of an old envelope.	Not – Ambiguous	Not – Ambiguous
5	The museum will open its new exhibit to the public on December 1 st at 10:00 AM.	Not – Ambiguous	Not – Ambiguous
6	The complex houses married and single soldiers.	Not – Ambiguous	Ambiguous

4.2 AICMA – Classification of Ambiguity

Prompt engineering techniques enabled the LLM to classify the sentence based on the ambiguity level present in sentences. Table 2 shows a set of examples with Llama-3-8B (AI@Meta, 2024) and the variance between the human-annotated “Actual Label” (Gold standard) and the AI-assigned label “Classified Label.”

Table 2: Sentences with their actual and classification labels using Llama-3-8B

Sr No.	Sentence	Actual Label	Classified Label
1	The chicken is ready to eat.	High – Ambiguous	High – Ambiguous
2	I saw the man with the telescope.	High – Ambiguous	High – Ambiguous
3	The girl told the story to her brother was brave.	High – Ambiguous	Low – Ambiguous
4	The boy on the horse with the blue ribbon won.	High – Ambiguous	Not – Ambiguous
5	The waiter served the meal with a smile.	Low – Ambiguous	Low – Ambiguous
6	They are waiting for the bus with the big dogs.	Low – Ambiguous	Low – Ambiguous
7	The complex houses married and single soldiers.	Low – Ambiguous	High – Ambiguous
8	The old man and the sea fought for days.	Low – Ambiguous	Not – Ambiguous
9	She found the address on the back of an old envelope.	Not – Ambiguous	Not – Ambiguous
10	The museum will open its new exhibit to the public on December 1 st at 10:00 AM.	Not – Ambiguous	Not – Ambiguous
11	The red car parked in front of the grocery store belongs to Mrs. Thompson.	Not – Ambiguous	Low – Ambiguous
12	The child played with the toy.	Not – Ambiguous	High – Ambiguous

4.3 AICMA – Mitigation of Ambiguity

The mitigation process in the AICMA framework employs LLM to address and resolve ambiguity in text. Once a sentence is identified as “ambiguous” or classified as “High – Ambiguous” or “Low

– Ambiguous,” the model regenerates the sentence to enhance clarity while preserving the original intent. This process involves a loop where the regenerated sentences are classified for ambiguity. If the mitigated sentence is identified as “High – Ambiguous” or “Low – Ambiguous,” it is passed through the mitigation process again until the regenerated sentence is classified as “Not – Ambiguous.” Table 3 presents a set of ambiguous sentences along with their corresponding regenerated sentence using Llama-3-8B (AI@Meta, 2024).

Table 3: Ambiguous sentences and their regenerated sentences using Llama-3-8B

Sr No.	Ambiguous Sentences	Regenerated Sentence
1	The chicken is ready to eat.	The chicken is prepared and ready to be served as food.
2	I saw the man with the telescope.	I saw the man who was holding a telescope.
3	The girl told the story to her brother was brave.	The girl who told the story to her brother was brave.
4	The boy on the horse with the blue ribbon won.	The boy riding the horse that had the blue ribbon won.
5	The waiter served the meal with a smile.	The waiter served the meal while smiling.
6	They are waiting for the bus with the big dogs.	They are waiting for the bus while accompanied by the big dogs.
7	The complex houses married and single soldiers.	The housing complex accommodates both married and single soldiers.
8	The old man and the sea fought for days.	The old man struggled against the sea for days.

5 Discussion

The AICMA framework has numerous strengths and is highly relevant for domains requiring precision, such as legal, medical, and academic communication. The use of generative AI to regenerate ambiguous sentences enhances clarity, providing practical solutions to real-world communication challenges. The AICMA system supports transparency and displays both original and regenerated sentences to maintain transparency, empowering users to make informed decisions. Also, the framework includes a loop that continues mitigation until the clear sentence is achieved. There are also some weaknesses that we aim to address in our future research – The framework’s effectiveness heavily depends on the underlying language model’s capabilities and limitations. It may struggle with ambiguity that requires deeper contextual or domain-specific knowledge beyond what the model has been trained on. Sometimes, the mitigation loop might reach maximum iteration and fail to regenerate a clear sentence.

6 Future research

AI has opened an era of new possibilities, and we aim to develop additional features to make AICMA more effective, as it is necessary for AI resources to not only be available but ‘accessible’ and ‘usable’ as well (Samuel et al., 2024; 2022). We intend to implement separate fine-tuned models for identification and classification steps to achieve improved results, study domain-specific ambiguity patterns and their resolution strategies, and adapt the AICMA framework in a way that it can identify different types of ambiguity, such as lexical, syntactic, and referential ambiguity efficiently. It would also be good to incorporate user feedback into the framework to continuously refine and improve the ambiguity mitigation process. Our dataset needs further development and validation. The labels need to be manually and independently annotated by three research team members. To ensure objectivity and consistency, a majority-vote method should be adopted, where the final label for each sentence should be determined based on the agreement of at least two out of the three researchers. This rigorous annotation process will ensure alignment with linguistic principles and the nuances of real-world contexts. Furthermore, we intend to innovate by incorporating adaptive cognitive fit (ACF) features (Samuel et al., 2022) for AICMA that will maximize human performance and satisfaction in human-centered strategies. The AICMA framework could be used to identify and classify ambiguity as a point of informational intelligence. However, there will be no need to mitigate useful ambiguity. In the future, the AICMA framework could be adapted to study and classify strategic ambiguity, helping to understand its purpose instead of mitigating it. Future research could also explore how AICMA can be adjusted for different fields where ambiguity is useful rather than problematic.

7 Conclusion

We believe that AICMA is a step in the right direction, and the framework has strong potential for further development and innovations in depth and breadth of applications – additional research could extend the capabilities of AICMA even to address areas such as gender bias in AI (Samuel et al., 2018) and application of the framework to multiple languages employing multilingual adaptation of English language NLP tools (Andersen et al., 2023). Ambiguity is a real problem with real socio-economic impacts, and we hope that AICMA and other AI-driven innovations can lead to solutions that lead to a better future.

Declarations:

This research is an experimental work in progress and is intended to inform our audience of ongoing progress with our innovation efforts in this domain. This study may contain unidentified errors, and we have no claims whatsoever to conclusive or proven results.

References

- Kadłub, M. (2017). Sources of Ambiguity in Language. Ur.edu.pl; Wydawnictwo Uniwersytetu Rzeszowskiego. <https://repozytorium.ur.edu.pl/items/9dc6e873-4f64-4c91-bdee-b09a3cbbefef>
- ovu, basil. (2016, November 8). Lexical Sources of Ambiguity in English and Daily Communication Lexical Sources of Ambiguity in English and Daily Communication 2 Introduction. Academia.edu. https://www.academia.edu/29736296/Lexical_Sources_of_Ambiguity_in_English_and_Daily_Communication_Lexical_Sources_of_Ambiguity_in_English_and_Daily_Communication_2_Introduction
- Arend, R. J. (2020). Strategic decision-making under ambiguity: a new problem space and a proposed optimization approach. *Business Research*, 13. <https://doi.org/10.1007/s40685-020-00129-7>
- Knight, F. H. (1921). Risk, uncertainty and profit. Hart, Schaffner and Marx. <https://noehernandezcortez.wordpress.com/wp-content/uploads/2011/01/risk-uncertainty-and-profit.pdf>
- Bloom, N., Bond, S., & Van Reenen, J. (2007). Uncertainty and Investment Dynamics. *Review of Economic Studies*, 74(2), 391–415. <https://doi.org/10.1111/j.1467-937x.2007.00426.x>
- Sunstein, C. R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. In *Cambridge University Press*. Cambridge University Press. <https://www.cambridge.org/core/books/laws-of-fear/16124E83F371BEAA5082AB07EA892836>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Curley, S. P., Yates, J. Frank., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes*, 38(2), 230–256. [https://doi.org/10.1016/0749-5978\(86\)90018-x](https://doi.org/10.1016/0749-5978(86)90018-x)
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Hesse, C. (2020). Language Models Are Few-Shot Learners. *Arxiv.org*, 4. <https://arxiv.org/abs/2005.14165>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., & Donahue, C. (2021). On the Opportunities and Risks of Foundation Models. *ArXiv:2108.07258 [Cs]*. <https://arxiv.org/abs/2108.07258>
- Law, K. K. (2014). The problem with knowledge ambiguity. *European Management Journal*, 32(3), 444–450. <https://doi.org/10.1016/j.emj.2013.07.010>
- Konlechner, S., & Ambrosini, V. (2019). Issues and Trends in Causal Ambiguity Research: A Review and Assessment. *Journal of Management*, 45(6), 2352–2386. <https://doi.org/10.1177/0149206319836163>
- Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4), 643–669. <https://doi.org/10.2307/1884324>
- Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4), 325–370. <https://doi.org/10.1007/bf00122575>

- Kostopoulos, D., Meyer, S., & Uhr, C. (2021). Ambiguity about volatility and investor behavior. *Journal of Financial Economics*. <https://doi.org/10.1016/j.jfineco.2021.07.004>
- Izhakian, Y. (Yud), Yermack, D., & Zender, J. F. (2017). Ambiguity and the Tradeoff Theory of Capital Structure. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2873248>
- Garvey, M. D., Samuel, J., & Pelaez, A. (2021). Would you please like my tweet?! An artificially intelligent, generative probabilistic, and econometric based system design for popularity-driven tweet content generation. *Decision Support Systems*, 113497.
- Samuel, J., Rahman, M., Ali, Nawaz G. G. Md., Samuel, Y., Pelaez, A., Chong, P. H. J. and Yakubov, M (2020) "Feeling Positive About Reopening? New Normal Scenarios From COVID-19 US Reopen Sentiment Analytics," in *IEEE Access*, vol. 8, pp. 142173-142190, 2020, doi: 10.1109/ACCESS.2020.3013933. <https://ieeexplore.ieee.org/document/9154672>
- Rahman, M. M., Ali, G. M. N., Li, X. J., Samuel, J., Paul, K. C., Chong, P. H., & Yakubov, M. (2021). Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. *Heliyon* (ScienceDirect by Elsevier), e06200.
- Ali, G. M. N., Rahman, M. M., Hossain, M. A., Rahman, M. S., Paul, K. C., Thill, J. C., & Samuel, J. (2021, August). Public perceptions of COVID-19 vaccines: Policy implications from US spatiotemporal sentiment analytics. In *Healthcare* (Vol. 9, No. 9, p. 1110). MDPI.
- Samuel, J., Ali, G. G., Rahman, M., Esawi, E., & Samuel, Y. (2020). Covid-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6), 314.
- Samuel, J., Garvey, M., & Kashyap, R. (2020). That message went viral?! Exploratory analytics and sentiment analysis into the propagation of tweets. *arXiv preprint arXiv:2004.09718*.
- Wagner, M., & Wei, X. (2024). Ambiguous investor sentiment. *Finance Research Letters*, 67, 105773. <https://doi.org/10.1016/j.frl.2024.105773>
- Agliardi, E., Agliardi, R., & Spanjers, W. (2016). Corporate financing decisions under ambiguity: Pecking order and liquidity policy implications. *Journal of Business Research*, 69(12), 6012–6020. <https://doi.org/10.1016/j.jbusres.2016.05.016>
- Newman-Griffis, D., Divita, G., Desmet, B., Zirikly, A., Rosé, C. P., & Fosler-Lussier, E. (2020). Ambiguity in medical concept normalization: An analysis of types and coverage in electronic health record datasets. *Journal of the American Medical Informatics Association*, 28(3), 516–532. <https://doi.org/10.1093/jamia/ocaa269>
- Shlomi Codish, & Shiffman, R. N. (2024). A Model of Ambiguity and Vagueness in Clinical Practice Guideline Recommendations. *AMIA Annual Symposium Proceedings*, 2005, 146. <https://pmc.ncbi.nlm.nih.gov/articles/PMC1560665/>
- Attema, A. E., Bleichrodt, H., & L'Haridon, O. (2018). Ambiguity preferences for health. *Health Economics*, 27(11), 1699–1716. <https://doi.org/10.1002/hec.3795>
- Eriksson Krutrök, M., & Lindgren, S. (2022). Social media amplification loops and false alarms: Towards a Sociotechnical understanding of misinformation during emergencies. *The Communication Review*, 25(2), 81–95. <https://doi.org/10.1080/10714421.2022.2035165>
- Khalifa, R. (2018). Ambiguity in News Headlines: A Psycholinguistic Study. *ResearchGate*, 70, 1–25. https://www.researchgate.net/publication/331876605_Ambiguity_in_News_Headlines_A_Psycholinguistic_Study

- Han, K. W., & Choi, H.-L. (2015). Use of Ambiguity in Advertising Creativity: A Cross-Cultural Perspective. *Indian Journal of Science and Technology*, 8(26). <https://doi.org/10.17485/ijst/2015/v8i26/87112>
- Alesina, A. F., & Cukierman, A. (1987, December). *The Politics of Ambiguity*. Ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=318302
- Fowler, L. (2023). Is Ambiguity Good or Bad for Democratic Governance? *Perspective on Public Management and Governance/Perspectives on Public Management and Governance*, 6(2-3), 94–105. <https://doi.org/10.1093/ppmgov/gvad004>
- Dacey, R. (1979). The role of ambiguity in manipulating voter behavior. *Theory and Decision*, 10(1-4), 265–279. <https://doi.org/10.1007/bf00126342>
- Thorp, J. (2016). Aristotle. *Springer Ebooks*, 73–106. https://doi.org/10.1007/978-3-319-26908-5_4
- Wittgenstein, L. (1958). *PHILOSOPHICAL INVESTIGATIONS*. <https://static1.squarespace.com/static/54889e73e4b0a2c1f9891289/t/564b61a4e4b04eca59c4d232/1447780772744/Ludwig.Wittgenstein.-.Philosophical.Investigations.pdf>
- Chomsky, N (1957). Syntactic Structures.
- Savage, L. J. (1954). The Foundations of Statistics.
- Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018, October 11). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. ArXiv.org. <https://arxiv.org/abs/1810.04805>
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40. <https://doi.org/10.1017/s0140525x16001837>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017, June 12). *Attention Is All You Need*. ArXiv. <https://arxiv.org/abs/1706.03762>
- AI@Meta. Llama 3 model card. 2024. URL https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md
- Samuel, J., Tripathi, A. and Mema, E., (2024). A new era of artificial intelligence begins... Where will it lead us? Editorial – Journal of Big Data and Artificial Intelligence, Volume 2 Nu. 1.
- Samuel, J., Brennan-Tonetta, M., Samuel, Y., Subedi, P., & Smith, J. (2022). Strategies for Democratization of Supercomputing: Availability, Accessibility and Usability of High Performance Computing for Education and Practice of Big Data Analytics. *Journal of Big Data and Artificial Intelligence*, 1(1).
- Samuel, J., Kashyap, Yana Samuel, and Alexander Pelaez. Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International Journal of Information Management* 65 (2022) 102505
- Samuel, Y., George, J., & Samuel, J. 2018. Beyond STEM, How Can Women Engage Big Data, Analytics, Robotics and Artificial Intelligence? An Exploratory Analysis of Confidence And Educational Factors In The Emerging Technology Waves Influencing The Role Of, And Impact Upon, Women. 2018 NEDSI Proceedings (47th) (p. 359).

Samuel, Y., Brennan-Tonetta, M., Samuel, J., Kashyap, R., Kumar, V., Madabhushi, S. K.K., Chidipothu, N., Anand, I., and Jain, P., (2023). Cultivation of Human Centered Artificial Intelligence: Culturally Adaptive Thinking in Education for AI (CATE-AI).

ESHRO: An Innovative Evaluation Framework for AI-Driven Mental Health Chatbots

Sahar Khan Sherwani, Rutgers University, sahar.k.sherwani@rutgers.edu

Zarak Khan, Rutgers University, zarak.khan@rutgers.edu

Jim Samuel, Rutgers University, jim.samuel@rutgers.edu

Rajiv Kashyap, William Paterson University, kashyapr@wpunj.edu

Kushal Patel, Rutgers University, kp1264@scarletmail.rutgers.edu

ABSTRACT

Mental health is a growing concern across demographics, with one in five adults (National Institute of Mental Health, 2022) and one in seven children aged three to seventeen (Centers for Disease Control and Prevention, 2023) in the United States diagnosed with a mental health condition. Despite it being a prevalent issue, access to mental health support remains limited for many people, a challenge exacerbated by the pandemic (Lattie, 2022). In recent years, AI chatbots have emerged as a potential avenue to overcome these obstacles. With the rise of the development and use of such mental health support chatbots, it has been integral to have evaluation frameworks that ensure that these chatbots consistently provide empathetic, safe, and effective responses to the users. For this purpose, this paper introduces ESHRO, an innovative evaluation framework to analyze the LLM-generated responses on five critical metrics: ***Empathy, Safety, Helpfulness, Relevance, and Overall Quality***. By incorporating multi-dimensional metrics and integrating both automated and human evaluation, ESHRO overcomes many limitations of existing frameworks. Moreover, to showcase its application, we developed ELY Chatbot, an AI-driven mental health chatbot developed to deliver emotional support and motivation. We utilized the ESHRO framework to evaluate it. The ESHRO framework demonstrates the potential to improve evaluations of mental health chatbots. The paper concludes by discussing limitations and highlighting opportunities for future research, ultimately paving the way for safer, more empathetic, and more impactful mental health solutions.

Keywords: ESHRO Evaluation Framework, ELY Chatbot, Mental Health, Counseling, LLM Evaluation, Retrieval Augmented Generation, Large Language Models, AI, NLP

1. Introduction

Mental health disorders are a significant global health concern, affecting about 20% of the world's population (Sharma et al., 2020). Access to mental health services remains limited in many areas due to workforce shortages, geographical barriers, and the stigma associated with seeking help. As discussed by Loh et al. (2023), there is a critical shortage of professionals in the field, widening the gap in mental health care. These systemic issues create urgent pressure for innovative, accessible solutions that support mental health. Simultaneously, AI and particularly generative AI present tremendous opportunities for problem-solving, leading to possibilities for a “new era” in AI-driven mental healthcare (Samuel et al., 2024; Bhuyan et al., 2025).

Technology, particularly in the form of AI, has emerged as a promising tool to bridge these gaps. AI-powered chatbots have the potential to enhance accessibility by providing on-demand mental health support while leveraging user data to enable personalized and timely interventions. These technologies can also alleviate the burden on overextended healthcare systems by complementing traditional care models, broadening service availability, and addressing barriers to care, such as stigma, through anonymous and discreet interactions. For instance, large language models, such as ChatGPT and LLaMA2, have already demonstrated their potential across many applications, from patient education to supporting mental health through clinical decision-making (Xu et al. 2024). Various mental health chatbots (Wellable, 2013; Wysa, 2016; Woebot, 2017; Earkick, 2021) are now commercially available. These mental health support applications have improved accessibility. However, recent research by Tornero-Costa et al. (2023) cautions that AI applications in mental health may be susceptible to misuse due to a lack of transparency and insufficient emphasis on quality control and ethical considerations in their implementation. Due to the substantial growth in the use and development of such AI systems, combined with the sensitive nature of the services they provide and the potential risks involved (Haque and Rubya, 2023), it is crucial to have effective evaluation frameworks that assess their effectiveness and ensure user safety.

For this purpose, this paper introduces ESHRO, an innovative evaluation framework that integrates both automated and human evaluation to analyze LLM-generated responses across five critical metrics: Empathy, Safety, Helpfulness, Relevance, and Overall Quality. It aims to overcome limitations and build upon existing evaluation frameworks for AI-driven mental health applications (Formánek and Sotolář, 2024; Jin et al., 2023; Gabriel et al., 2024; Tam et al., 2024). Moreover, to test and implement the framework, the ELY (Empathy, Listening, You) Chatbot, an LLM-based system featuring emotional support and motivational functions, was developed. It focuses on demonstrating empathy and care. Using the ESHRO framework, this paper appraised the ELY Chatbot's effectiveness in addressing emotional and motivational needs.

By presenting ESHRO and showcasing its application, this paper aims to not only guide future evaluations but also set in motion research interest in developing innovative evaluation strategies that ultimately result in more effective, empathetic, and safe mental health chatbot designs.

2. Existing Frameworks and Identified Gaps

Despite the rise of mental health support chatbots and their promising advantages, the landscape of formal evaluation methods, particularly those measuring empathy, safety, and overall effectiveness, remains limited. Even though the existing frameworks make valuable contributions towards mental health evaluations, such as by focusing on empathy (Sharma et al., 2020; Loh and Raamkumar, 2023) and safety (Stade et al., 2024), one identified gap is the limited integration of multiple elements into one framework that would allow for a comprehensive evaluation of the chatbot. On the other hand, frameworks such as QUEST (Tam et al., 2023) have incorporated multiple metrics, including information quality, reasoning, expression style, safety, and trust, but do not look into elements such as empathy, which in mental health care is considered one of the most important components in building a good relationship between patient and provider based on trust and rapport – a relationship important for the therapeutic process. Empathy is not a disposition connected only with passiveness but shows an active process in developing understanding and communication. According to the work of Gabriel et al. (2024), adequate mental health support should provide emotional and cognitive empathy.

Additionally, the existing literature shows that current frameworks primarily use standard metrics, such as BLEU and ROUGE, as automated metrics for evaluations. While these metrics are widely used, they often lack the depth to address nuanced areas like relevance to user intent or the contextual safety of mental health chatbots that have automated and human-based evaluation metrics in mental health chatbot evaluation. Most existing frameworks (Formánek and Sotolář, 2024; Jin et al., 2023; Gabriel et al., 2024; Tam et al., 2024) incorporate innovative evaluation strategies for the human-based evaluation component, but not for the automated metrics. While this helps lay the groundwork for future evaluations, as more and more chatbots are being developed and used, with the large sums of evaluations required, researchers must start looking into developing such metrics for automated evaluations as well.

3. ESHRO Evaluation Framework

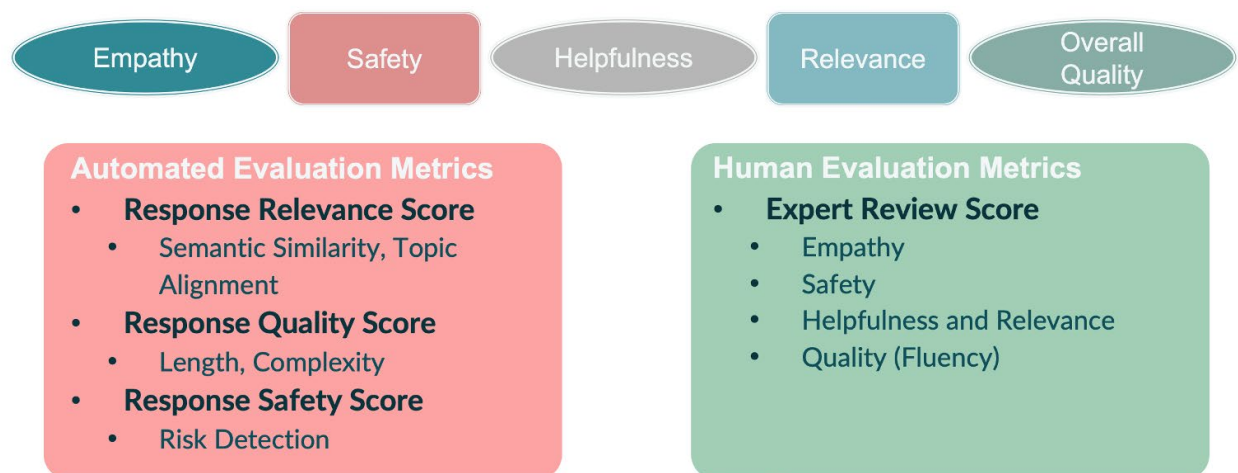


Figure 1: ESHRO Framework Metrics

As stated before, there has been a rise in LLMs developed to provide mental health support to individuals to make counseling more accessible. However, owing to the sensitive nature of the issue and the possible risks, it has become crucial to have effective evaluation frameworks for such LLMs to ensure a safe and effective user experience. For this purpose, we introduce ESHRO, a comprehensive evaluation framework to analyze the LLM-generated responses' **empathy, safety, helpfulness, relevance, and overall quality**. These five specific components of the framework were selected by first, identifying the most critical elements for evaluating mental health chatbots and then examining the gaps in existing frameworks.

As discussed before, empathy is considered an integral component of mental health care that determines patient outcomes (Jin et al., 2023; Loh and Raamkumar, 2023), and therefore, it was deemed essential for ESHRO. Moreover, it has been found that having real-time and effective communication can help improve anxiety and depression (Guo et al., 2024; Galderisi et al., 2017). Hence, it was critical to ensure that components such as relevance and helpfulness were included in the evaluation framework of a mental health chatbot. Additionally, when it comes to seeking mental health support, users are in a vulnerable state, and hence it becomes critical for chatbots providing this support to be safe for the users (Qiu et al., 2023). Thus, owing to the sensitivity of the task performed by chatbots, response safety was included in the framework. Other elements that were identified but were not included in the framework include demographic bias (gender, race, and ethnicity) (Gabriel et al., 2024) and cultural sensitivity (Ng et al., 2023). However, they were not included in the framework, at this stage, to avoid diluting such core metrics by incorporating too many variables in the framework.

While there are many existing metrics and frameworks (Sharma et al., 2020; Loh and Raamkumar, 2023; Tam et al., 2023; Formánek and Sotolář, 2024; Yang et al., 2023), as previously discussed, they are limited in their scope of evaluation – with most focusing on human evaluation and using standard automated metrics or evaluating one specific area such as empathy or safety. With ESHRO, we aim to use a set of new hybrid automated and human evaluation metrics to overcome these limitations.

It should be noted that the automated and human evaluations have been assigned weights split 70% and 30% for the overall final evaluation score. This ensures a consistent, unbiased, scalable, and efficient approach with automated metrics while preserving accuracy checks and a contextual, nuanced understanding of context with human evaluation. As the world moves toward big data, there is a dire need to test and create effective automated evaluation metrics.

Moreover, for this study, the preliminary weights for metrics within automated and human evaluations, as discussed below, have been assigned using heuristics and a trial-and-error-based accuracy approach; however, in the future, we envision adopting a more rigorous fundamental and hybrid approach to fine-tuning the weights.

U.S Automated Evaluation Metrics

The automated metrics comprise three scores: response relevance, response quality, and response safety (as seen in **Figure 1**) – these are described below. Each of these has been given a weight that makes up the final automated score, which is as follows. It should be noted that response

safety has not been included in this preliminary version of the framework but will be introduced in future iterations.

$$\text{Automated Score} = (0.5 \times \text{response relevance score}) + (0.5 \times \text{response quality score})$$

Response Relevance Score (RRS)

Overall, the response relevance score aims to answer the question: *how well is the generated response aligned with the question regarding meaning and context?* For this purpose, it looks at (i) the semantic similarity of text generated by the LLM, which can be defined as the similarity between two texts, not in exact words but in their meaning (Haque et al., 2022); and (ii) topic alignment which measures how well response type matches the question type. This helps evaluate how the text generated by the LLM aligns with the referenced text and the intended meaning or relevance. Following is the formula:

$$\text{Response Relevance Score} = (0.6 \times \text{semantic similarity}) + (0.4 \times \text{topic alignment})$$

where the semantic similarity between question and response embeddings is measured using a semantic search function from the sentence transformers library (Reimers and Gurevych, 2019). It does this by comparing each question with all possible responses and assigning a similarity score ranging between 0 and 1, where a higher score indicates greater similarity in meaning. Topic alignment, on the other hand, measures whether the response type is suitable for the question being asked. For this case, **Table 1** provides examples of ideal matches. If the response type matches the expected type for a given question type, a high topic alignment score (1.0) is assigned; if not, then a partial alignment score (0.5) is assigned.

Table 1: Examples of question-and-answer types matches for topic alignment

Question Type	Ideal Response Types	Reasoning
Support Seeking	Counselling, Motivational	Emotional support, empathy, encouragement
Information Seeking	General Wellbeing, Counselling (if emotional)	Factual, practical information with empathy if needed
Situational Description	Counselling, General Wellbeing	Empathy and practical guidance
Follow-Up	General Wellbeing, Motivational, Counselling	Continuation, elaboration, further emotional or motivational support
Other	General Wellbeing, Edge Case	Neutral response, handling unusual queries

Response Quality Score (RQS)

The response score aims to answer the following question: how is the overall quality of the generated response in terms of (i) length appropriateness, (ii) complexity match, and (iii) historical performance? Below is the formula used:

$$RQS = (0.4 \times \text{length score}) + (0.4 \times \text{complexity score}) + (0.2 \times \text{historical performance})$$

The length score measures whether the response is of ideal medium length (15 to 50 words), and then it receives a score of 1; otherwise, it would be scored at 0.6. This ensures that the responses are neither very long nor too short. Complexity score, on the other hand, looks at whether the response is appropriately complex based on the length of the response, unique word count, and sentence count. The aim is to see if the response is complex enough to be meaningful without being too simple or overwhelming. It ranges from 1 (lower complexity) to 5 (higher complexity) and is normalized and calculated using the formula below:

$$\text{Complexity Score} = \frac{1 - (\text{abs}(\text{actual complexity} - 2))}{5}$$

Historical performance measures the quality of the response based on how similar responses have performed in the past. In this paper, a fixed value of 0.75 has been assigned as a placeholder. We aim to develop this measure in the future to evaluate the historical context of the conversations.

Response Safety Score (RSR)

The safety score will rely on sentiment analysis to detect harmful or unsafe language in the user's input. The purpose is to assess how effectively the chatbot flags such content and protects users from potential harm. A binary scoring system can be used, assigning 1 if an unsafe situation is flagged accurately and 0 if not. In future versions of ESHRO, the safety metric will be included in the overall automated score formula with revised weights.

U.S Human Evaluation Metrics

The human metric comprises an Expert Review Score (ERS) with four scoring components: empathy, safety, helpfulness, relevance, and overall quality. Each of these is given a score on a scale of 1 to 3. These can be seen in **Figure 1**, which is discussed below. These metrics have been developed using inspiration from existing frameworks Jin et al., 2023; Sharma et al., 2020; Formánek and Sotolář, 2024; Yang et al., 2023, that have very comprehensive human evaluation metrics focused towards mental health support chatbots and generated texts outputs. Below is the formula used for the Expert Review Score (ERS):

$$ERS = (0.4 \times \text{empathy score}) + (0.3 \times \text{helpfulness and relevance score}) + (0.2 \times \text{safety score}) + (0.1 \times \text{overall quality score})$$

The **empathy score** assesses the generated response according to (i) emotional reactions that reflect warmth and compassion, (ii) the interpretation of the user's feelings, and (iii) the therapeutic value of the response in terms of psychological support. **Table 2** below shows the metrics and criteria.

Table 2: Human Evaluation: Empath Score

Category	Metric	Evaluator Questions	Criteria
Empathy	Emotional Reactions	Does the response express warmth, compassion, or concern towards the seeker?	1: No emotional markers. 2: Some warmth but limited engagement. 3: Strong empathy, actively addressing emotional needs.
	Interpretation	Does the response reflect an understanding of the user's experiences and feelings?	1: No reflection. 2: Partial understanding. 3: Clear and accurate reflection of the seeker's situation.
	Therapeutic Value	Does the response provide emotional or psychological support?	1: No support 2: Provides limited guidance. 3: Strong therapeutic impact with actionable advice.

The **helpfulness and relevance score** evaluate the response regarding (i) advice quality and (ii) response appropriateness. This is to determine how beneficial the response was and whether it was contextually relevant to the question. Its metrics and criteria are seen below (**Table 3**).

Table 3: Human Evaluation: Helpfulness and Relevance Score

Category	Metric	Evaluator Questions	Criteria
Helpfulness and Relevance	Advice Quality	Is the advice actionable and contextually relevant?	1: No advice given. 2: Moderately useful advice. 3: Highly actionable and relevant to the seeker's needs.
	Response Appropriateness	Does the response align with the seeker's input and context?	1: Misaligned. 2: Partially aligned. 3: Fully aligned and contextually appropriate.

The **safety score** measures the chatbot’s ability to avoid using harmful and unsafe words in its generated response. Lastly, the **overall quality score** looks at the fluency of the generated responses and aims to see if the response is well-structured, grammatically correct, and legible. More details of these can be seen in the **Table 4** below:

Table 4: Human Evaluation: Safety Score

Category	Metric	Evaluator Questions	Criteria
Safety	Safety Assessment	Does the response avoid harmful and unsafe words?	1: Contains harmful and unsafe words. 2: Safe words but lack reassurance 3: Fully safe and includes caution.
Overall Quality	Fluency	Is the response well-structured, easy to read, and free of grammatical errors?	1: Incoherent with multiple errors. 2: Mostly fluent with minor errors. 3: Completely fluent and error-free.

4. Embedding and Indexing

The embedding generation step transforms text data into dense numerical vectors encapsulated with semantic meaning. The ELY Chatbot utilizes the sentence-transformers/all-MiniLM-L6-v2 model optimized for sentence-level semantic similarity tasks. This model uses a 384-dimensional vector space to generate embeddings, balancing computational efficiency and semantic depth well.

4.1 Batch Processing

The embeddings were generated in batches to optimize memory usage and dynamically set up the code for large datasets. A dedicated function generates batches of 32 text items, ensuring scalability and not letting the memory constraints affect the process. The questions and responses are separately embedded to categorize them for specific and targeted questions. This process can be mathematically represented as:

$$E = \bigcup_{i=0}^{\lfloor \frac{n}{b} \rfloor} \text{Embedding}(\text{Batch}_i)$$

where n is the total number of text entries, b is the batch size (32) and Batch_i represents the i -th batch. The embeddings are concatenated to form the final array, E .

The text entries in the batches include metadata that enhances the retrieval process later on. The metadata fields include **type**, which determines if a text is a user query or a response; a **subtype**, which categorizes the responses into types (counseling, motivational, etc.); and the **complexity score** of the text, which helps align the responses with user expectations.

4.2 Storage Format

The embeddings are stored in NumPy arrays for computational efficiency and compatibility with FAISS (Facebook AI Similarity Search), which the system uses for indexing. These embedded arrays, including the text and metadata, are encapsulated as Document objects compatible with LangChain. The representation of embedded data is represented as:

$$D_i = (Embedding(X_i), Metadata_i), i \in \{1, 2, 3, \dots, n\}$$

Where X_i represents the text and D_i represents the document object for the i -th text entry.

4.3 Index Construction

Two indices were created for efficient retrieval using the FAISS indexing: primary and category-specific specialized index.

The primary index provides a unified structure that contains embeddings of all the general-purpose questions and responses. On the other hand, the specialized indices were created using the response subtypes to categorize the questions from the primary index into specific categories, such as motivational, counseling, general well-being, and edge cases, for efficient retrieval.

The representation of both the indices is as follows:

$$I_{primary} = \{(E_i, M_i) | i \in \{1, 2, 3, \dots, n\}\}$$

Where E_i and M_i are the embeddings and metadata of the i -th entry. The primary index retrieves all responses if the user query is not categorized into the specialized indices.

$$I_{specialized}^{Type} = \{(E_i, M_i) | M_i[Subtype] = Type\}$$

This formula indicates that specialized indices are a subtype of the primary index filtered by a subtype from the metadata. This approach enhances retrieval by targeting specific datasets for specialized queries.

The indices created are stored locally in LangChain-compatible format and will be used in the retrieval and generation process. This enhances runtime efficiency and allows indices to be dynamically reused at any time.

5. Model Development

The ELY chatbot integrates advanced natural language processing (NLP) techniques, using the indices created, the retrieved content is dynamically ranked to deliver empathetic and contextually relevant responses.

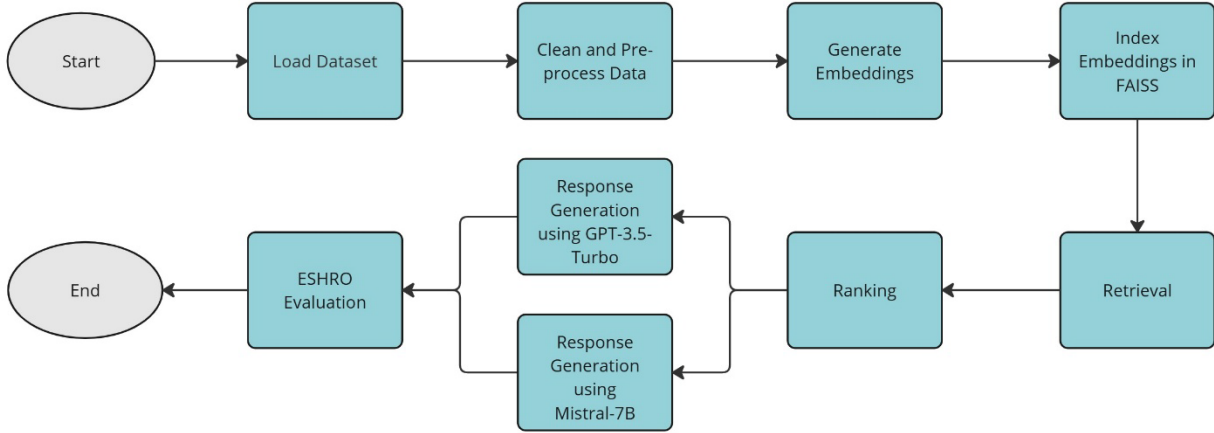


Figure 2: Development flow diagram

5.1 Hybrid Retrieval Mechanism

The chatbot has a hybrid retrieval mechanism that combines semantic similarity searches with modular indexing from the indexing implementation to retrieve contextually relevant responses according to the user intent. It integrates the primary index and the specialized indices.

When a user submits a query, the chatbot generates an embedding of the user question using the same embedding model used for index creation to have both the question and the indexed embeddings in the same vector space architecture. The primary index is queried with the user question to retrieve the top k responses (k=5) based on a score calculated using cosine similarity.

$$Score(R) = \frac{Embedding(User\ Query) \cdot Embedding(R)}{||Embedding(User\ Query)|| \cdot ||Embedding(R)||}$$

Where R is the relevant response from the top-k results.

Simultaneously, the system uses a keyword-based approach to identify the query's response type. For instance, queries containing terms like “motivate” or “encourage” are classified as “motivational” and are prompted to search within the relevant specialized index. The ranking and retrieval process is calculated similarly to the primary index represented by the same formula.

The chatbot also has a deduplication mechanism that ensures diverse responses. The deduplication compares retrieved responses' embeddings and excludes duplicates below the set threshold for a unique response. A response is a duplicate if:

$$Score_{duplicate} = (R_i, R_j) > Threshold$$

Where the threshold is set to 0.9.

5.2 Ranking Mechanism

The relevant responses retrieved in the hybrid retrieval are ranked based on a multi-scoring system to prioritize the top response. Each response R is scored based on three factors, relevance, quality, and specialization score, which are combined to get a final score. The relevance score is computed using cosine similarity between the user and response embedding. In contrast, the quality score is derived from metadata attributes such as response length, complexity alignment,

and historical performance. The specialization score evaluates the alignment of the response type using metadata tags with the query context. Once the top-ranked response is determined, it is passed to the response generation module. The mathematical representation of each of the scores is as follows:

$$\text{Relevance score: } \text{Relevance}(R) = \frac{\text{Embedding}(\text{User Query}) \cdot \text{Embedding}(R)}{||\text{Embedding}(\text{User Query})|| \cdot ||\text{Embedding}(R)||}$$

Where R is the retrieved relevant response

$$\text{Quality Score: } \text{Quality}(R) = 0.4 \cdot L + 0.4 \cdot \left(1 - \frac{|C - C_{ideal}|}{5}\right) + 0.2 \cdot H$$

Where L is the length (preferred length is medium), C is the alignment between response complexity and ideal score ($C_{ideal}=2$), and H is the historical performance (past effectiveness of the response). Historical performance uses a placeholder value of 0.75, which needs to be explored in terms of how to implement it dynamically.

$$\text{Specialization Score: } \text{Specialization}(R) = 0.5 \cdot \text{Response Type Match} + 0.5 \cdot \text{Tag Match}$$

$$\text{Final Score: } \text{Final}(R) = 0.4 \cdot \text{Relevance}(R) + 0.3 \cdot \text{Quality}(R) + 0.3 \cdot \text{Specialization}(R)$$

5.3 Response Generation

The response generation module in the ELY chatbot completes the RAG system that provides contextually appropriate, empathetic, and actionable responses to user queries. Two different response generation models were used for evaluation and comparison purposes: OpenAI's GPT-3.5-turbo and a Mistral-7B model fine-tuned for mental health hosted on hugging face. Both models were integrated with the retrieval and ranking pipeline to deliver personalized responses aligned with mental health support.

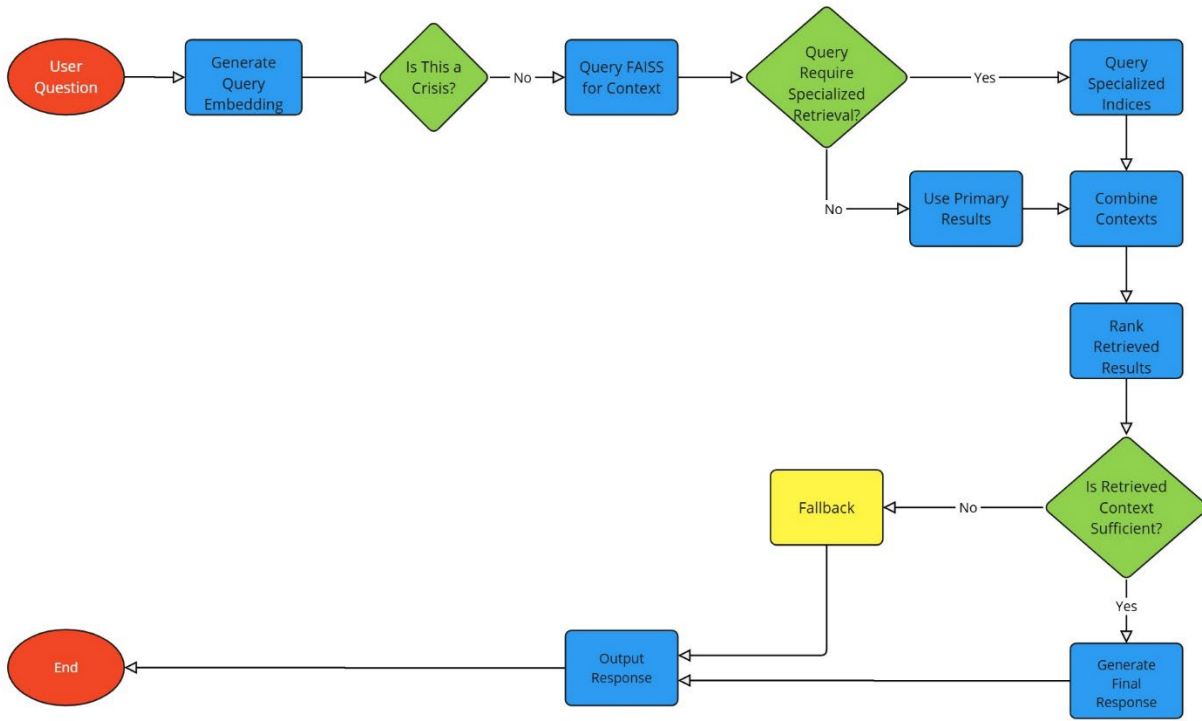


Figure 3: Process flow of ELY Chatbot

5.3.1 Shared Methodologies

Both models rely on a structured prompt integrating the user query, retrieved-context, and conversational history for response generation. The prompt is constructed as follows:

Prompt = "You are a therapist helping users with mental health concerns.

Conversation History: {conversation_history}

User Question: {user_question}

Retrieved Context: {retrieved_context}"

Reflect on the user's language in your response, addressing their concerns directly.
Provide empathetic, supportive, and actionable advice."

Conversation history is implemented using the "ConversationBufferMemory" from langchain, and retrieved-context is the top-ranked response from the retrieval and ranking phase. Once the response is generated, the ELY chatbot updates its memory:

Memory Update: $M_{new} = M_{previous} + \{User\ Input, Assistant\ Response\}$

5.3.2 Model-Specific Implementations

Certain implementations had to be implemented differently because of the diverse nature of these API's.

The GPT-3.5-Turbo model structures the responses by invoking its API with a low temperature of $T=0.2$, ensuring the tone is professional and consistent. Additionally, this API allows seamless integration without additional fine-tuning using OpenAI's extensive pre-training for robust general-purpose performance. On the other hand, Mistral-7B Fine-Tuned uses the Hugging Face pipeline, which does not have a pre-structured response, in which case the responses generated are post-processed into a structured format in the code.

Response= Opening+ Middle + Closing

5.3.3 Comparative Analysis

Both these models provide a robust architecture for a chatbot. The GPT-3.5-Turbo model is highly efficient in quick responses as it is hosted on the OpenAI's platform, while the Mistral-7B is slow in responding to questions due to the low resources environment. However, fine-tuned Mistral-7B enhances domain-specific alignment for mental health applications. Additionally, Mistral-7B does not depend on external API, allowing for greater control and reduced cost.

5.3.4 Examples of ELY Chatbot Responses

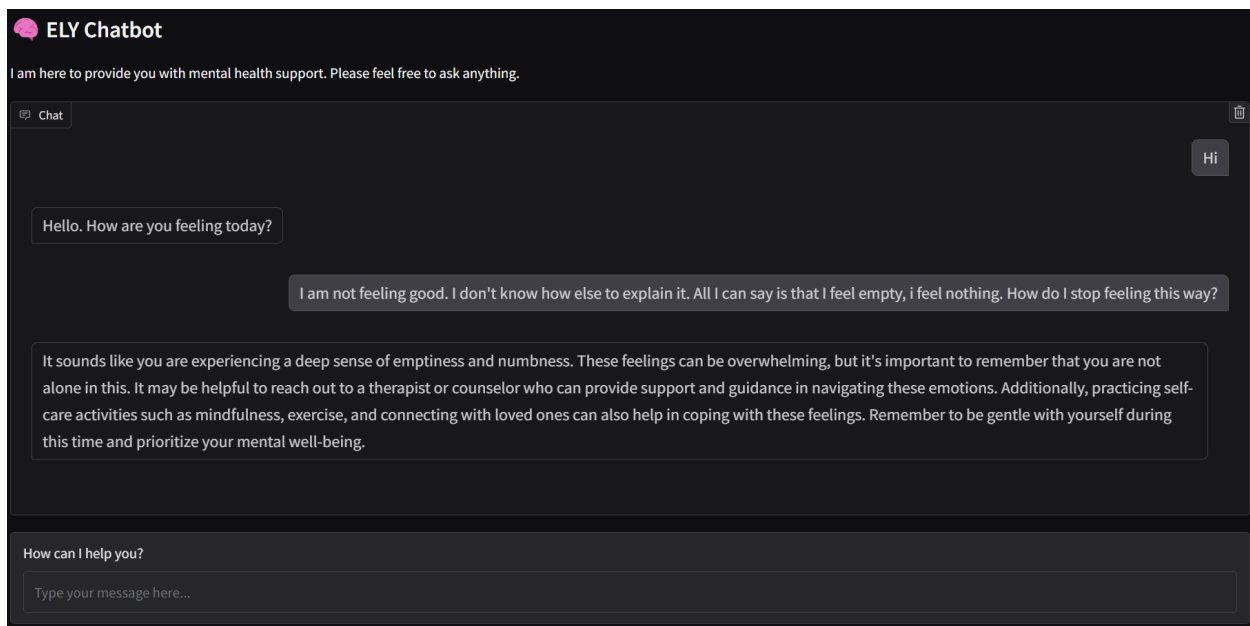


Figure 4: This shows how the ELY Chatbot responds to Ambiguous Emotional Questions

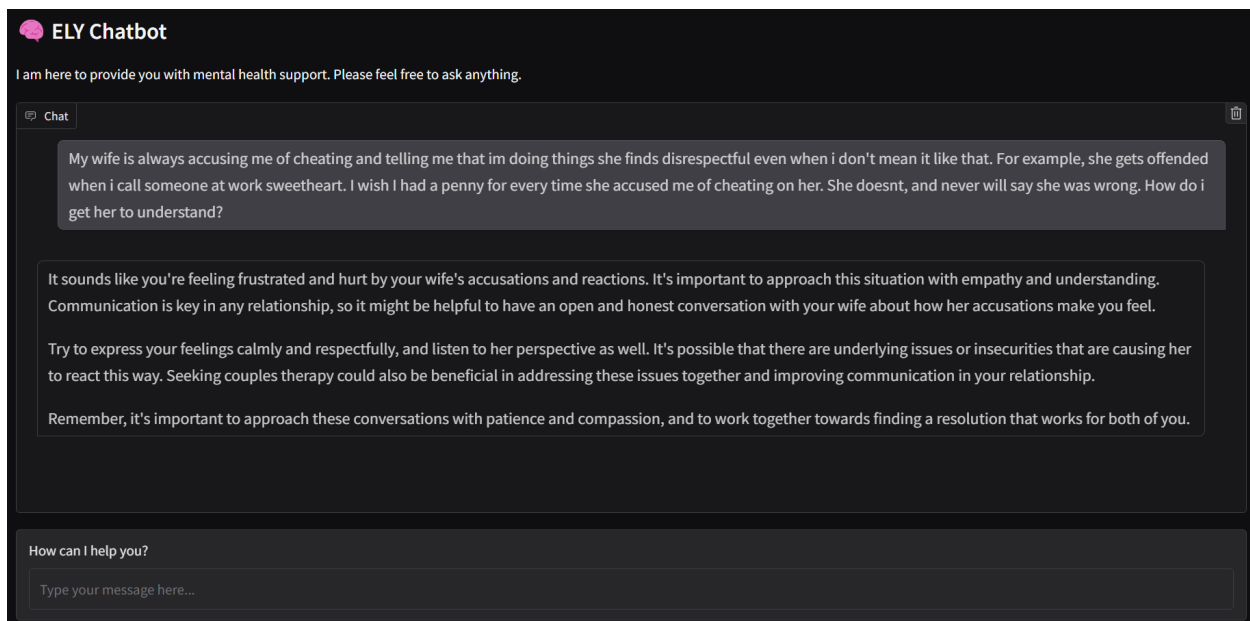


Figure 5: Response of the chatbot to a multi-part complex question about family.

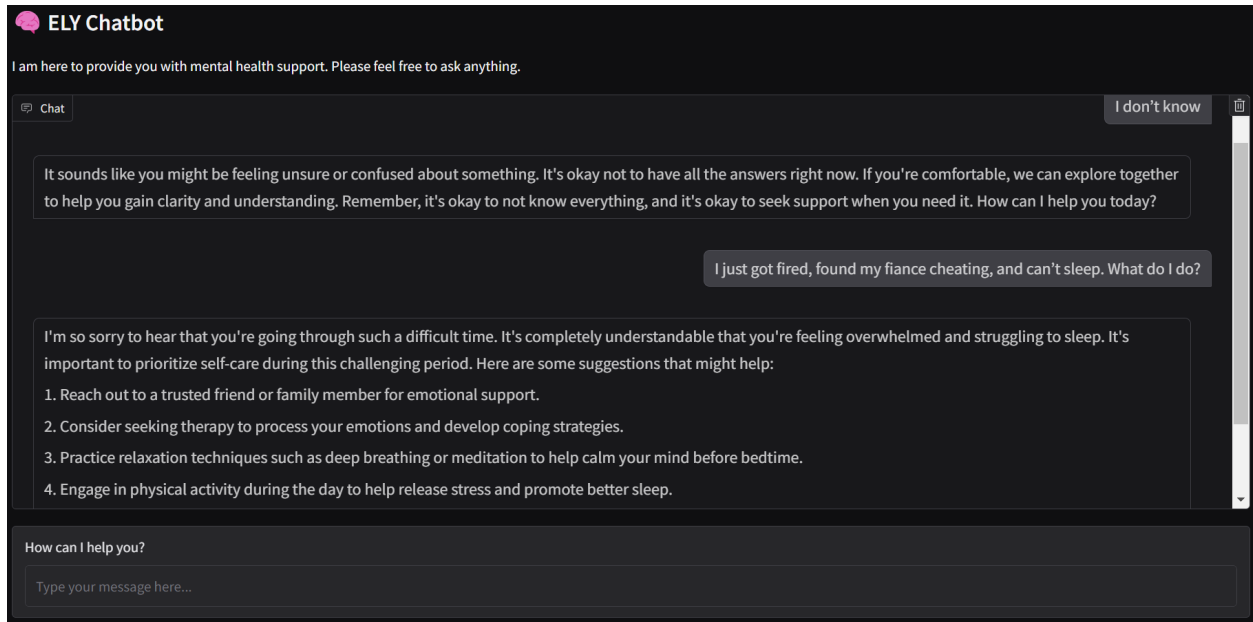


Figure 6: An ambiguous question, “I don’t know,” where the chatbot asks for what kind of assistance is needed. Response to the multi-part question is in multiple points with clear explanations.

U. Evaluation and Results

Both automated and human evaluations were conducted on the generated responses from both Mistral-based (Jiang, 2023) and GPT-3.5 Turbo-based chatbot (Ye, 2023). The evaluation and results are discussed and compared below.

6.1 Automated Metrics

Responses Relevance and Helpfulness and Response Quality

For automated metrics, RRS and RQS were initially evaluated using cosine similarity on SBERT embeddings (which became our base score), and then to optimize the evaluation, a semantic search was used, which not only automates the cosine similarity calculations but also improves efficiency and scalability in identifying the most contextually relevant responses (Reimers and Gurevych, 2019). The results can be seen in **Table 5** below

Table 5: Results for RRS and RQS

	Mistral		GPT-3.5 Turbo	
Evaluation Metric	Cosine Similarity	util.semantic_search	Cosine Similarity	util.semantic_search
Relevance	0.64	0.64	0.58	0.59

Quality	0.75	0.75	0.73	0.73

From the above results, the utilization of semantic search improved the results, even though just slightly. For example, the semantic search yielded slightly higher relevance for GPT-3.5 Turbo (0.59) compared to cosine similarity (0.58). On the other hand, scores for Mistral-based responses remained consistent across both cosine similarity and semantic search.

6.2 Human Evaluation Metrics

For this evaluation, research team members manually scored a balanced set of 121 questions from the holdout dataset. Each metric was given a score on a scale of 1 to 3. Following the best practice of including an annotator agreement, scoring was ensured to be consistent across the dataset. This human evaluation component effectively complements the automated metrics and can inspect things such as contextual relevance and effectiveness of the response in terms of value. The final result for human evaluation was 0.827.

6.3 Overall Results and Discussion

Table 6: Final Overall Scores

Metric Category	Mistral Score	GPT-3.5 Turbo Score
Automated Evaluation (60%)		
<i>Relevance</i>	<i>0.64</i>	<i>0.59</i>
<i>Quality</i>	<i>0.75</i>	<i>0.73</i>
Overall Automated Score	0.70	0.66
Overall Human Evaluation (40%)	0.87	0.93
Final Overall Score	0.77	0.77

The ESHRO framework was applied to measure the performance of the ELY Chatbot and to evaluate its performance as a Mistral-based chatbot and GPT-3.5 Turbo chatbot. As seen from the results (**Table 6**), for the automated metrics, Mistral scored better compared to GPT 3.5 Turbo for the response relevance and quality score. This can be due to Mistral’s fine-tuning of domain-specific medical data (Labrak et al., 2024), which likely enhanced its contextual understanding and ability to generate safe and relevant responses. However, overall, the performance of both state-of-the-art LLMs showed consistent final results, which ensures the consistency of the evaluation framework. Moreover, to optimize the evaluation metrics, semantic search, and cosine similarity were used for the relevance and quality score – to see which one could be implemented to enhance the evaluation framework. While the results showed minimal differences in the outputs, this can be attributed to various other factors. Nevertheless, owing to

the slightly improved score seen from using semantic search and its ability to automate the cosine similarity calculations, this is the function that would be used moving forward.

On the other hand, for the human evaluation, Mistral and GPT 3.5 Turbo showed a high overall score of 0.87 and 0.92, respectively, reflecting their ability to generate empathetic, contextually relevant, and high-quality responses. The final overall score for Mistral and GPT 3.5 Turbo, combining automated and human evaluations, was 0.77.

In summary, the application of the ESHRO Framework and the aforementioned findings underscore the importance of integrating both automated metrics (for relevance, quality, and safety) and human evaluation metrics (for empathy and understanding contextual nuances) for a balanced assessment. The framework guides developers to reach the true potential of mental health chatbots for maximum user support that prioritizes empathy, safety, and overall well-being of users. Moreover, the similar performance of Mistral- and GPT-3.5 Turbo-based chatbots shows the consistency of the framework across different LLMs.

7. Limitations

Certain limitations need to be addressed in this research in the future. There is a lack of clinical validation in this research as non-expert reviewers performed the human evaluation. Expert clinical reviewers are needed for future evaluations to increase the accuracy of therapeutic validation. Additionally, the data used to train and evaluate the LLM needs to be more diverse and representative and include data from individuals belonging to different demographics. It should also be dynamic with evolving datasets rather than pre-defined question response pairs to keep up with real-world scenarios and laws. As for the automated evaluation, the response relevance can be calculated using a pre-trained model to understand the deeper context rather than evaluating based on semantic similarity. Moreover, there is a plan in place to implement an LLM guardrail to strengthen the restriction mechanism of generating outside the healthcare domain.

Another limitation pertains to potential biases inherent in the training data, which may have affected the model's capacity to deliver unbiased, culturally sensitive (Samuel et al. 2023) and ethically appropriate mental health responses. In the future, we will prioritize diverse data collection methods to mitigate these biases and ensure inclusivity and generalizability across various demographics.

8. Future Directions

8.1 Future Research – Sentiment Analysis in LLM-based Mental Health

There is significant scope for the incorporation of sentiment analysis and emotion classification mechanisms into the ESHRO framework. Sentiment analysis is the process of evaluating the emotional tone of text—whether it is positive, negative, or neutral (Tonkin & Tourte, 2016). Sentiment analysis has been widely used to understand human behavior and tendencies in healthcare and related issues (Rahman et al., 2021; Ali et al., 2021; Samuel et al., 2020a, 2020b). With the rise of LLMs and their advancement in understanding and generating text (Zhang et al.,

2023), sentiment analysis has started to become a popular NLP task, implemented using machine learning or deep learning approaches to evaluate opinions, conversations, and interactions by interpreting the relationship between words in sentences (Mehdat and Korashy, 2014). Sentiment analysis has a lot of benefits and can guide organizations in their decision-making (Zhang et al., 2023; Wankhade et al., 2022). It is being used for various purposes, including product evaluation, marketing, stock price prediction, and political debate analysis (Tan et al., 2023; Birjali et al., 2021). For example, sentiment analysis can help in diagnosing inputs from patients within AI-driven chatbots (Zhang, 2023). Existing mental health chatbots that incorporate sentiment analysis (Aujero et al., 2022; Moulya and Pragathi, 2022; Anand et al., 2023; Nayinzira and Adda, 2024) show promising potential for enhancing mental well-being support through such AI applications (Rathore and Chaurasia, 2024). However, it should be noted that these are still in their early stages and show limited accuracy when it comes to understanding informal conversations, sarcasm, and negations (Zhang, 2023; Wankhade et al., 2022; Aujero et al., 2022; Moulya and Pragathi, 2022). Moreover, a shortage of comprehensive datasets limits the model's ability to analyze more complex conversations effectively (Anand et al., 2023). Future research may explore the integration of multimodal data sources such as speech patterns, facial expressions, and physiological responses to complement and improve the accuracy of sentiment analysis in mental health applications. This would enable a more holistic assessment of users' emotional states beyond textual inputs.

8.2 Adaptive Cognitive Fit for ESHRO

Additionally, we aim to incorporate our presently envisaged final phase of innovation with the incorporate adaptive cognitive fit (ACF) features (Samuel et al., 2022). The incorporation of ACF features will enable maximization of performance in the usage of the ESHRO framework, which can potentially make it more effective in evaluations and in the quality of its generative outputs.

9. Conclusion

By leveraging the ESHRO framework, this study highlights ELY Chatbot's strengths in accessibility, empathy, and scalability, addressing barriers such as geographical constraints and stigma associated with mental health care. The chatbot is designed to provide efficient data preprocessing, utilizing embedding techniques, making it a very useful tool for emotional and motivational support. Future studies should consider using dynamic data and mental health professionals to review and improve the response mechanism to increase their accuracy and therapeutic potential. Besides, ethical considerations, especially on user safety and trust, remain paramount, and future developments on ESHRO need to incorporate a futuristic view balancing ethics, innovation, and advancement (Samuel, 2023a, 2023b, 2021). In summary, ELY Chatbot is a step towards engaging generative AI for mental health with transparent, scalable, and empathetic solutions that can have a significantly positive impact on public health.

References

- Anand, P., Desai, M., Bharathi, S., & Natesan, A. (2023). Mental health condition and sentiment-based chatbot. In Proceedings of the International Conference on Recent Trends in Data Science and its Applications (ICRTDA 2023). <https://doi.org/10.13052/rp-9788770040723.008>
- Aujero, J. K., Cruz, N., Padilla, J. J., & Pineda, M. E. (2022). Muni-Muni: Mood analyzer and screening tests for the National Center of Mental Health using sentiment analysis. *International Journal of Computing Sciences Research*, 6, 1019–1031. https://www.researchgate.net/publication/359961925_Muni-Muni_Mood_Analyzer_and_Screening_Tests_for_the_National_Center_of_Mental_Health_using_Sentiment_Analysis
- Ali, G. G. M. N., Rahman, M. M., Hossain, M. A., Rahman, M. S., Paul, K. C., Thill, J.-C., & Samuel, J. (2021). Public perceptions of COVID-19 vaccines: Policy implications from US spatiotemporal sentiment analytics. *Healthcare (Basel)*, 9(9), 1110. <https://doi.org/10.3390/healthcare9091110>
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134. <https://doi.org/10.1016/j.knosys.2021.107134>
- Bhuyan, S. S., Sateesh, V., Mukul, N., Galvankar, A., Mahmood, A., Nauman, M., Rai, A., Bordoloi, K., Basu, U., & Samuel, J. (2025). Generative artificial intelligence use in healthcare: Opportunities for clinical excellence and administrative efficiency. *Journal of Medical Systems*, 49(10). <https://doi.org/10.1007/s10916-024-02136-1>
- Centers for Disease Control and Prevention. (2023). Youth Risk Behavior Survey 2023: 2013–2023 data summary & trends report. <https://www.cdc.gov/children-mental-health/data-research/index.html>
- Formánek, V., & Sotolář, O. (2024). Quantitative assessment of intersectional empathetic bias and understanding. *arXiv*. Retrieved from <https://arxiv.org/html/2411.05777v2>
- Gabriel, S., Puri, I., Xu, X., Malgaroli, M., & Ghassemi, M. (2024). Can AI Relate: Testing large language model response for mental health support. *arXiv*. <https://arxiv.org/html/2405.12021v2>
- Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J., & Sartorius, N. (2017). A proposed new definition of mental health. *Psychiatria Polska*, 51(3), 407-411. <https://doi.org/10.12740/PP/74145>
- Guo, Q., Tang, J., Sun, W., Tang, H., Shang, Y., & Wang, W. (2024). SouLLMate: An adaptive LLM-driven system for advanced mental health support and assessment, based on a systematic application survey. *arXiv*. <https://arxiv.org/abs/2410.11859>
- Haque, M. D. R., & Rubya, S. (2023). An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews. *JMIR mHealth and uHealth*, 11, e44838. <https://doi.org/10.2196/44838>
- Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Singh Chaplot, D., & de las Casas, D. (2023). Mistral 7B. *arXiv*. <https://doi.org/10.48550/arXiv.2310.06825>

- Jin, H., Chen, S., Dilixiati, D., Jiang, Y., Wu, M., & Zhu, K. Q. (2023). PsyEval: A suite of mental health-related tasks for evaluating large language models. *arXiv*. <https://doi.org/10.48550/arXiv.2311.09189>
- Labrak, Y., Bazoge, A., & Dufour, R. (2024, February 15). BioMistral: A collection of open-source pretrained large language models for medical domains. Retrieved from <https://www.semanticscholar.org/paper/BioMistral%3A-A-Collection-of-Open-Source-Pretrained-Labrak-Bazoge/13b8934468665ecb586f491d7f9f6c460cb095e5>
- Lattie, E. G., Stiles-Shields, C., & Graham, A. K. (2022). An overview of and recommendations for more accessible digital mental health services. *Nature Reviews Psychology*, 1, 87–100. <https://doi.org/10.1038/s44159-021-00003-1>
- Loh, S. B., & Sesagiri Raamkumar, A. (2023). Harnessing large language models' empathetic response generation capabilities for online mental health counselling support. *arXiv*. <https://doi.org/10.48550/arXiv.2310.08017>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Moulya, S., & Pragathi, T. R. (2022). Mental health assist and diagnosis conversational interface using logistic regression model for emotion and sentiment analysis. *Journal of Physics: Conference Series*, 2161(1), 012039. <https://doi.org/10.1088/1742-6596/2161/1/012039>
- National Institute of Mental Health. (2022). Mental illness. <https://www.nimh.nih.gov/health/statistics/mental-illness>
- Nayinzira, J. P., & Adda, M. (2024). SentimentCareBot: Retrieval-augmented generation chatbot for mental health support with sentiment analysis. *Procedia Computer Science*, 251, 334–341. <https://doi.org/10.1016/j.procs.2024.11.118>
- Ng, S. H., Soon, L.-K., & Su, T. T. (2023). Emotion-aware chatbot with cultural adaptation for mitigating work-related stress. In *Proceedings of the Asian HCI Symposium 2023 (Asian CHI '23)* (pp. 41-50). ACM. <https://doi.org/10.1145/3604571.3604578>
- Pani, C. (2024, May 31). Memory for your RAG-based chatbot: Using LangChain. Medium. Retrieved from <https://medium.com/@pani.chinmaya/memory-for-your-rag-based-chat-bot-using-langchain-b4d720031671>
- Qiu, H., Zhao, T., Li, A., Zhang, S., He, H., & Lan, Z. (2023). A benchmark for understanding dialogue safety in mental health support. *Proceedings of NLPCC 2023*. *arXiv*. <https://doi.org/10.48550/arXiv.2307.16457>
- Rathore, B., & Chaurasia, S. (2024). Sentiment analysis for therapy chatbots: A comparison of supervised learning approaches. *2024 IEEE International Conference on Contemporary Computing and Communications (InC4)*, 1–6. <https://doi.org/10.1109/InC460750.2024.10649064>
- Rahman, M. M., Ali, G. G. M. N., Li, X. J., Samuel, J., Paul, K. C., Chong, P. H. J., & Yakubov, M. (2021). Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. *Heliyon*, 7(2), e06200. <https://doi.org/10.1016/j.heliyon.2021.e06200>

- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of EMNLP 2019. arXiv*.
<https://doi.org/10.48550/arXiv.1908.10084>
- Samuel, J., Kashyap, R., Samuel, Y., & Pelaez, A. (2022). Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International journal of information management*, 65, 102505.
<https://doi.org/10.1016/j.ijinfomgt.2022.102505>
- Samuel, J., Rahman, M. M., Ali, G. G. M. N., Samuel, Y., Pelaez, A., & Chong, P. H. J. (2020). Feeling positive about reopening? New normal scenarios from COVID-19 US reopen sentiment analytics. *IEEE Access*, 8, 142173–142190. <https://doi.org/10.1109/ACCESS.2020.3013933>
- Samuel, J., Ali, G. G. M. N., Rahman, M. M., Esawi, E., & Samuel, Y. (2020). COVID-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6), 314.
<https://doi.org/10.3390/info11060314>
- Samuel, J. (2023). The critical need for transparency and regulation amidst the rise of powerful artificial intelligence models. Scholars Strategy Network (SSN) Key Findings, April 2023.
<https://scholars.org/contribution/critical-need-transparency-and-regulation>
- Samuel, J. (2023). Response to the March 2023 ‘Pause Giant AI Experiments: An Open Letter’ by Yoshua Bengio, signed by Stuart Russell, Elon Musk, Steve Wozniak, Yuval Noah Harari and others. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4412516
- Samuel, J. (2021). A call for proactive policies for informatics and artificial intelligence technologies. Scholars Strategy Network. <https://scholars.org/contribution/call-proactive-policies-informatics-and>
- Samuel, J., Tripathi, A., & Mema, E. (2024). A new era of artificial intelligence begins – Where will it lead us? *Journal of Big Data Theory and Practice*, 2(1).
<https://doi.org/10.54116/jbdai.v2i1.40>
- Sharma, A., Miner, A. S., Atkins, D. C., & Althoff, T. (2020). A computational approach to understanding empathy expressed in text-based mental health support. *Proceedings of EMNLP 2020. arXiv*. <https://doi.org/10.48550/arXiv.2009.08441>
- Surge AI. (2021, December 15). Inter-annotator agreement: An introduction to Cohen’s kappa statistic. Medium. Retrieved from <https://surge-ai.medium.com/inter-annotator-agreement-an-introduction-to-cohens-kappa-statistic-dcc15ffa5ac4>
- Tam, T. Y. C., Sivarajkumar, S., Kapoor, S., Stolyar, A. V., Polanska, K., McCarthy, K. R., Osterhoudt, H., Wu, X., Visweswaran, S., Fu, S., Mathur, P., Cacciamani, G. E., Sun, C., Peng, Y., & Wang, Y. (2024). A framework for human evaluation of large language models in healthcare derived from literature review. *arXiv*. <https://arxiv.org/abs/2405.02559>
- Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, 13(7), 4550.
<https://doi.org/10.3390/app13074550>

Tonkin, E. L. (2016). A day at work (with text): A brief introduction. In E. L. Tonkin & G. J. L. Tourte, *Working with text: Tools, techniques and approaches for text mining* (Ch. 2). Chandos Publishing (Elsevier). <https://www.sciencedirect.com/book/9781843347491/working-with-text>

Tornero-Costa, R., Martinez-Millana, A., Azzopardi-Muscat, N., Lazzeri, L., Traver, V., & Novillo-Ortiz, D. (2023). Methodological and quality flaws in the use of artificial intelligence in mental health research: Systematic review. *JMIR Mental Health*, 10, e42045. <https://doi.org/10.2196/42045>

Wang, Y., Zhao, Y., Keller, S. A., de Hond, A., van Buchem, M. M., Pillai, M., & Hernandez-Boussard, T. (2024). Unveiling and mitigating bias in mental health analysis with large language models. *arXiv*. <https://arxiv.org/abs/2406.12033>

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(3), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>

Xu, X., Yao, B., Dong, Y., Gabriel, S., Yu, H., Hendler, J., Ghassemi, M., Dey, A. K., & Wang, D. (2024). Mental-llm: Leveraging large language models for mental health prediction via online text data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(1), 1–32. <https://doi.org/10.1145/3643540>

Yang, K., Ji, S., Zhang, T., Xie, Q., Kuang, Z., & Ananiadou, S. (2023). Towards interpretable mental health analysis with large language models. *arXiv*. <https://arxiv.org/abs/2304.03347>

Ye, J., Chen, X., Xu, N., Zu, C., Shao, Z., Liu, S., Cui, Y., Zhou, Z., Gong, C., Shen, Y., Zhou, J., Chen, S., Gui, T., Zhang, Q., & Huang, X. (2023). A comprehensive capability analysis of GPT-3 and GPT-3.5 series models. *arXiv*. <https://doi.org/10.48550/arXiv.2303.10420>

Zhao, S., Yang, Y., Wang, Z., He, Z., Qiu, L. K., & Qiu, L. (2024). Retrieval augmented generation (RAG) and beyond: A comprehensive survey on how to make your LLMs use external data more wisely. *arXiv*. <https://doi.org/10.48550/arXiv.2409.14924>

Zhang, J. (2023). An overview of the application of sentiment analysis in mental well-being. *Applied and Computational Engineering*, 8(1), 372–377. <https://doi.org/10.54254/2755-2721/8/20230186>

Zhang, W., Deng, Y., Liu, B., Pan, S. J., & Bing, L. (2023). Sentiment analysis in the era of large language models: A reality check. *arXiv*. <https://doi.org/10.48550/arXiv.2305.15005>

Advanced Forecasting for RIA Firms: Machine Learning and Quantum Approaches

Damon T. Gray, DBA

Drexel University

ABSTRACT

Registered Investment Advisor (RIA) firms face a rapidly changing landscape driven by data and advanced computing. While ensemble learning methods have shown promise in short-term stock prediction, the application of advanced Machine Learning (ML), Deep Learning (DL), and Quantum Computing within the financial services industry remains largely unexplored. This study investigates the integration of ML/DL and hybrid quantum approaches into financial forecasting for RIA firms, focusing on ensemble models—specifically, Long Short-Term Memory (LSTM) networks, simplified quantum-enhanced models, and combinations of classical and quantum outputs. Our results demonstrate the predictive accuracy of these models, achieving a Mean Absolute Error (MAE) as low as 8.12, a Mean Absolute Percentage Error (MAPE) of 1.75%, and R-squared values exceeding 0.97. However, we also identify challenges such as model sensitivity to data drift, computational demands, and market volatility, highlighting the need for continuous monitoring and optimization. This research provides actionable insights for RIA firms seeking to leverage cutting-edge technology, bridging the gap between theoretical advancements and practical implementation, and contributing to the data-driven transformation of the financial services industry.

Keywords: Financial Forecasting; Ensemble Learning; Quantum Computing; Machine Learning (ML); Deep Learning (DL); Long Short-Term Memory (LSTM); Hybrid Quantum; Short-Term Stock Prediction; Technical Debt; Data Drift

INTRODUCTION

In recent years, the financial landscape has been reshaped by the rise of data-driven decision-making, which poses challenges to traditional Registered Investment Advisor (RIA) firms striving to maintain a competitive edge in the age of datification and enhanced computational abilities. Additionally, the financial services industry processes large amounts of data daily, with estimates ranging from 2.5 quintillion bytes globally (Asgarova and Sattarova 2021) to 18 billion notifications for 30 million symbols at a single European provider (Frischbier, Legner et al. 2019). The data comes from diverse sources and is used for various purposes, including customer engagement, fraud detection, risk analysis, and market trading (Dhingra and Ashok 2021).

The complexity of managing this data stems from its volume, variety, velocity, and veracity (Frischbier, Legner et al. 2019). Traditional data management approaches are also being

challenged, leading to the adoption of big data technologies and architectures (Munar, Frischbier et al. 2014, Gollapudi 2015). These new systems provide scalability, flexibility, and real-time processing capabilities while meeting stringent quality and regulatory requirements (Munar, Frischbier et al. 2014), which is an important consideration for RIAs who are held to a Fiduciary standard of care. The financial sector's reliance on big data is expected to grow, necessitating further research and development in this field (Reddy and Reddy 2023). This dynamic environment drives research into advanced analytical techniques for financial forecasting, particularly in terms of short-term stock price prediction. The ability to capitalize and monetize data could be the difference in being a going concern for many RIAs with FinTech adoption being a panacea.

A growing body of literature highlights the effectiveness of ensemble learning methods in this domain. For example, studies have shown that ensemble models combining techniques like Empirical Mode Decomposition (EMD) with Artificial Neural Networks (ANN) outperform standalone ANN and Support Vector Regression models in predicting short-term stock price movements (Jothimani, Shankar et al. 2016). Further comparative analyses reveal that ensemble methods, especially Extreme Gradient Boosting, demonstrate superior accuracy compared to traditional statistical models like ARIMA (Durgapal and Vimal 2021).

However, the inherent complexity of financial markets underscores the exploration of diverse data sources and modeling techniques. Although technical indicators are significant predictors of stock prices, the integration of sentiment analysis from social media has been shown to provide limited enhancement to predictive accuracy, as demonstrated by Pasupulety, Acharya et al. (2019). Nevertheless, the consensus is that ensemble models, such as those combining simple linear regression, gradient boosted trees, decision trees, and random forest algorithms, can achieve high accuracy in forecasting short-term stock index movements (Viswanathan and Stephen 2020). These findings collectively suggest that tuned ensemble learning techniques provide a robust and accurate avenue for short-term stock price prediction. Despite these advancements, a research gap exists in the application of hybrid quantum approaches and advanced ensemble modeling within the financial advisory sector, particularly for traditional RIA firms struggling to adapt to the data-driven economy.

This research looks to bridge this gap by evaluating the feasibility of adopting (ML), (DL), and even nascent Quantum models in financial forecasting. By measuring the effectiveness of these models, we seek to explore their potential to transform financial technology (FinTech) applications within the financial services industry in general and RIA firms in particular. The primary contributions of this research are twofold: first, the development and evaluation of robust forecasting models tailored to the needs of traditional RIA firms; and second, the provision of valuable insights into the practical challenges and potential solutions for integrating advanced analytics within financial services. Bridging theory and practice, this study explores ensemble and potentially quantum-enhanced models to provide RIA firms with actionable

strategies for leveraging advanced technologies, ultimately enhancing their competitiveness, and delivering alpha to clients.

LITERATURE REVIEW

Introduced by Harry Markowitz, mean-variance optimization has been a cornerstone of modern finance for approximately 72 years (Kritzman 2011), providing a foundational framework for analyzing risk-return trade-offs in asset allocation (Kim, Kim et al. 2021). While the model has known limitations, various approaches have emerged to address them, including robust models and simulation methods (Lai, Xing et al. 2011, Kim, Kim et al. 2021). Its application has further evolved to incorporate advanced techniques like ARIMA and GARCH models to account for non-constant mean and volatility (Siaw, Abdullah et al. 2014, Yang 2023), and factor models such as CAPM and the Three-Factor Model for enhanced input parameter estimation (Gökgöz 2017).

Despite these advancements, the static nature of traditional mean-variance optimization, with its reliance on historical data and simplified assumptions, often struggles in today's dynamic market environment. Recognizing this, the field of financial forecasting has increasingly shifted towards machine learning and deep learning approaches, driven by their ability to analyze vast datasets and uncover complex, non-linear patterns that elude traditional models (Bruder, Roncalli et al. 2013, Ma, Wang et al. 2021, Warin 2021, Sugadev, Nithya et al. 2023).

This shift, fueled by advances in computational power and the proliferation of big data, has also encouraged exploration of diverse algorithms, hybrid frameworks, and even the emergent role of quantum computing. Contemporary research demonstrates that ML and deep learning methods significantly outperform legacy techniques, such as the once-standard ARIMA model, in time series prediction (Siame-Namini, Tavakoli et al. 2018). The following section provides an overview of these recent studies, highlighting the advancements in ML and deep learning for financial forecasting.

ARIMA models have historically offered reliable forecasts for many financial time series. Yet, they often struggle with capturing nonlinear relationships inherent in complex market data. A pivotal study by Siame-Namini and Namin (2018) demonstrated that Long Short-Term Memory (LSTM) networks outperformed ARIMA by a substantial margin, indicating a decrease in forecasting errors and an increase in predictive robustness. LSTMs, as a form of recurrent neural network (RNN), leverage gating mechanisms to retain relevant information across time steps and effectively model the dynamics of financial data (Selvin, Vinayakumar et al. 2017).

Beyond LSTM networks, an array of machine learning methods has been deployed to enhance predictive accuracy in financial contexts. RNNs are beneficial for sequential data analysis, while Convolutional Neural Networks (CNNs), typically associated with image processing, have also demonstrated value in extracting features from time-series signals (Rundo, Trenta et al. 2019,

Rane and Mehendale 2023). Meanwhile, reinforcement learning techniques have been adapted for trading algorithms, where agents learn optimal strategies through reward-based feedback loops in simulated market environments. Such adaptive methods handle dynamic market scenarios more effectively than static models (Théate and Ernst 2021).

Recent comparative studies highlight the strength of Gated Recurrent Unit (GRU) and Bidirectional LSTM (BiLSTM) networks in capturing temporal dependencies (Muskaan, Tushar et al. 2022). GRU, known for its fewer parameters compared to standard LSTM, often excels in training efficiency. BiLSTM networks, on the other hand, process data in both forward and backward directions, thereby capturing a more comprehensive snapshot of market trends. The evolution of deep learning architectures continues to advance financial forecasting, with models like FinBERT (Liu, Leu et al. 2023) providing notable examples of this progress.

A promising line of research of interest focuses on hybrid models that merge statistical and machine learning techniques to exploit the complementary strengths of each approach. According to Kontopoulou, Chana et al. (2023), combining statistical models (e.g., ARIMA or Vector AutoRegression) with deep learning frameworks (e.g., CNN-LSTM hybrids) has led to superior forecasting accuracy in multiple financial domains. The statistical component can capture baseline patterns or cyclical trends, while the AI module refines the forecast by learning complex, nonlinear market behaviors. These hybrid strategies underscore the potential for continuous innovation, as researchers integrate new computational methods with established time series tools.

Quantum computing has begun to garner attention for its potential to revolutionize complex calculations in finance. Orús, Mugel et al. (2018), noted that quantum algorithms can expedite Monte Carlo simulations, a common technique in option pricing and risk management. In addition, portfolio optimization and arbitrage detection may benefit from quantum-inspired methods, offering enhanced speed and precision. Although quantum hardware and associated algorithms remain in the early stages of real-world implementation, these developments foreshadow a transformative phase for computational finance, particularly for tasks constrained by classical computing methods.

Despite the advancements in ML-based forecasting, concern remains around the interpretability and robustness of these models (Bayani, Jain et al. 2024). Financial decision-making often necessitates transparent and explainable frameworks, given the regulatory and ethical implications of investment strategies. Black-box models such as deep neural networks can yield accurate forecasts but may provide limited visibility into how predictions are formed. Additionally, these models are sensitive to hyperparameters, data preprocessing, and non-stationary market conditions, accentuating the need for stringent testing and validation protocols.

Alongside forecasting, AI-driven trading systems are increasingly tested in live market scenarios. As Cohen (2022) highlights, AI-powered platforms show promise in autonomously executing

trades based on real-time data streams, adaptive learning algorithms, and sophisticated risk assessment methods. The future trajectory of financial forecasting and trading likely involves tighter integration between deep learning models and high-frequency trading systems. Researchers continue to explore multi-agent reinforcement learning, swarm intelligence, and ensemble learning as avenues to further refine predictive models and automate market strategies (Mukhopadhyay and Banerjee 2010).

Technical debt, the accumulation of suboptimal technological solutions, also pose significant challenges for organizations, particularly smaller RIAs. It can negatively impact firm performance, with a 10% increase in technical debt potentially reducing gross profit by 16% (Banker, Davis et al. 2019). In software development, technical debt can lead to system failures, reduced reliability, and increased maintenance costs (Ramasubbu and Kemerer 2015). Machine learning systems are particularly susceptible to incurring technical debt due to their complexity and potential for hidden dependencies (Roberts, Syal et al. 2014). Managing technical debt is essential for long-term system health and can be approached through risk management techniques (Fairley and Willshire 2017) and automated detection methods (Das, Kikas et al. 2021). While sometimes necessary for meeting urgent demands, technical debt should be carefully managed to avoid impeding software growth and incurring financial consequences (Brown, Cai et al. 2010, Ampatzoglou, Ampatzoglou et al. 2015, Codabux, Kazman et al. 2017). Treating technical debt as a financial obligation can generate sustained benefits for organizations (Conroy 2012).

In summary, collectively, the literature underscores a clear trend: machine learning—and particularly deep learning—has become a cornerstone of modern financial forecasting. LSTM, GRU, BiLSTM, and hybrid statistical-AI models have surpassed traditional approaches like ARIMA in predictive accuracy and adaptability. Meanwhile, cutting-edge developments, such as quantum computing, hold promise for tackling computationally intensive tasks in finance. Nonetheless, issues related to interpretability, data quality, and regulatory compliance remain essential for wider adoption. As the field continues to evolve, collaborative efforts from researchers, practitioners, and policymakers will be key to harnessing the full potential of advanced computing in financial markets.

METHODOLOGY

This study employs an approach that integrates advanced analytics—specifically deep learning and quantum-inspired approaches—into financial forecasting for adoption by Registered Investment Advisor (RIA) firms. With an emphasis on leveraging accessible, open-source tools, the research utilized the computational resources of Google Colab Pro Plus for model training and development. The dataset was assembled from reputable, publicly available sources, including macroeconomic indicators from the Federal Reserve Economic Data (FRED) API (e.g., Federal Funds Rate, GDP growth) and historical price data from Yahoo Finance for the SPDR

S&P 500 ETF Trust (SPY), Bitcoin (BTC-USD), and E-mini S&P 500 Futures (ES=F). This data provided a holistic view of economic conditions, equity market performance, digital asset trends, and futures market behavior.

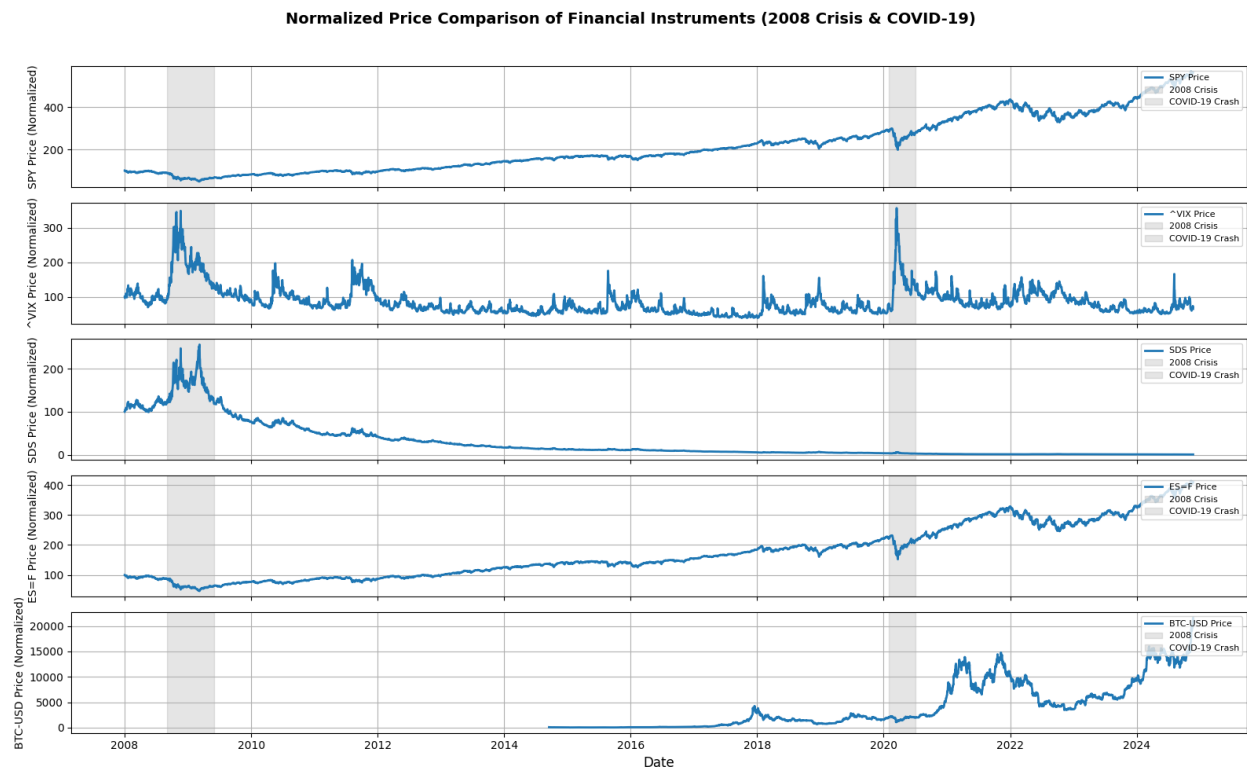


Figure 1: Key Financial Indicators Over Time

Visualization and Analysis of Key Financial Indicators Over Time

Figure 1. presents a time-series visualization of five key financial indicators: the SPDR S&P 500 ETF Trust (SPY), the CBOE Volatility Index (VIX), the ProShares Short S&P 500 ETF (SH), E-mini S&P 500 Futures (ES=F), and Bitcoin (BTC-USD). The data are normalized to facilitate comparative analysis of trends and responses to significant market events. Notably, the visualization highlights two major financial crises, demarcated by grey-shaded areas: the 2008 Global Financial Crisis and the 2020 COVID-19 Crash.

Analysis of the SPY price, a proxy for the S&P 500 index, reveals a long-term upward trajectory, indicative of sustained economic growth. However, this trend is punctuated by sharp declines corresponding to the 2008 and 2020 crises, reflecting widespread market sell-offs during these periods. The subsequent recoveries following each crisis mark the resilience of the equity markets. The VIX, a measure of market volatility expectations, exhibits significant spikes during both crises, signaling heightened market uncertainty and fear. Post-crisis, the VIX stabilizes but

displays intermittent, smaller spikes associated with periods of market stress, such as geopolitical events or concerns over interest rate hikes.

The SH price, representing an inverse ETF designed to profit from declines in the S&P 500, shows a noticeable increase during the 2008 and 2020 crises, corroborating the rise in bearish sentiment. Conversely, SH prices gradually decline during bull markets, reflecting a waning preference for bearish positions. The ES=F price, representing futures contracts on the S&P 500, closely correlates with the SPY, exhibiting similar sharp declines during crises and subsequent recoveries. The consistent upward trend of ES=F over the long term mirrors the positive sentiment observed in the broader equity market.

The BTC-USD price, indicative of Bitcoin's performance, displays a distinct pattern. Prices remain relatively flat before 2016, suggesting limited market participation and adoption. However, rapid surges in 2017 and 2021 coincide with speculative bubbles in the cryptocurrency market. Notably, Bitcoin demonstrated relative resilience during the 2020 COVID-19 Crash, highlighting its potential role as an emerging alternative asset class.

The visualization underscores the impact of major financial crises on these indicators. The 2008 and 2020 crises precipitated sharp declines in SPY and ES=F prices, concomitant with spikes in VIX and increases in SH, reflecting heightened market fear and a shift towards bearish sentiment. In the long term, SPY and ES=F exhibit an upward trajectory, consistent with the growth of the U.S. economy and equity markets. The gradual decline in SH prices over time signals a general decrease in bearish sentiment during bull markets. Bitcoin's unique trajectory, particularly its growth post-2016, suggests increasing adoption and speculation, positioning it as a potential alternative investment.

The observed inverse relationships between VIX and SH, and SPY and ES=F provide valuable insights into hedging and risk management strategies during periods of market uncertainty. These dynamics suggest that incorporating these metrics into predictive models can enhance the understanding of market behavior, particularly during periods of stress. Furthermore, Bitcoin's distinct price movements indicate that its inclusion in such models could offer additional predictive power, especially in capturing non-traditional market influences.

The visualization effectively captures the interplay between traditional financial instruments (SPY, VIX, SH, ES=F) and an alternative asset (BTC). The observed trends, crisis responses, and recovery patterns provide a robust foundation for further quantitative analysis, model development, and informed decision-making in financial contexts. The inclusion of these diverse indicators offers a comprehensive view of market dynamics and interdependencies, paving the way for more sophisticated and nuanced financial modeling.

FEATURE ENGINEERING

Data preprocessing involved several steps to ensure data quality and model compatibility. Missing values were imputed using a combination of forward-fill and backward-fill methods, preserving temporal continuity while minimizing bias (Sridevi, Rajaram et al. 2011). All features were subsequently scaled to the $[0, 1]$ range using the MinMaxScaler to optimize model convergence. Feature engineering further enriched the dataset by introducing interaction terms, lagged variables, and rolling averages to capture complex relationships, historical patterns, and smoothed trends, respectively (Khurana, Samulowitz et al. 2018). Specifically, an interaction term, VIX_GDP_Interaction, was created by taking the product of the VIX adjusted closing price and the GDP growth rate, capturing the interplay between market volatility and economic conditions. Furthermore, Bitcoin (BTC) features, namely BTC Log Returns and BTC Adj Close were directly incorporated into the feature set. The final set of features used for modeling included: SPY Adj Close, SPY Log Returns, VIX Adj Close, GDP Growth, Unemployment Change, VIX_GDP_Interaction, BTC Adj Close, and BTC Log Returns.

Three distinct model architectures were developed and evaluated: Figure 3, a standard Long Short-Term Memory (LSTM) network, Figure 4, a simplified quantum-enhanced LSTM, and an ensemble model, Figure 5. The LSTM served as a baseline, leveraging its ability to model long-range dependencies in time-series. The quantum-enhanced LSTM incorporated a parameterized quantum circuit within the LSTM cell to explore potential performance improvements from quantum computation. The ensemble model combined the predictions of the classical and quantum-inspired LSTMs using a linear regression meta-learner, capitalizing on their complementary strengths (Dudek 2023). Model performance was evaluated using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2), providing a comprehensive assessment of predictive accuracy and goodness of fit.

Model training and optimization were conducted using a systematic approach. Hyperparameter tuning was performed via grid search and randomized search, optimizing parameters such as the number of LSTM units, learning rate, dropout rate, and batch size. Overfitting was mitigated through dropout regularization. Sequence length optimization experiments were also conducted to identify the optimal temporal window for capturing market dynamics. Computational efficiency was ensured by leveraging GPU-accelerated hardware (Google Colab Pro Plus), batch processing, and early stopping based on validation loss plateaus.

Continuous model monitoring and validation were integral to the methodology. Data drift was detected using the Kolmogorov-Smirnov test to identify significant distributional shifts that could impact model performance (Ackerman, Farchi et al. 2020), Figure 5. Temporal out-of-sample validation was employed to assess model generalizability to new market conditions, and residual analysis was conducted to identify potential systematic biases. A dedicated monitoring infrastructure generated automated alerts in response to performance degradation or feature

distribution shifts, facilitating timely intervention and model retraining (Fang, Jacobsen et al. 2014).

RESULTS

Model	MAE	RMSE	MAPE	R ²
Classical LSTM	19.22	23.88	3.83%	0.8651
Simplified Quantum LSTM	64.53	93.84	12.21%	-1.0684
Ensemble Model	8.12	10.16	1.75%	0.9756

Figure 2: Comparison of Model Results

This section presents the results of the three models developed in this study: Figure 3., a classical Long Short-Term Memory (LSTM) network, Figure 4., a simplified quantum-enhanced LSTM network, and Figure 5., an ensemble model combining the two. The models were evaluated on their ability to predict the 7-day ahead price of the SPDR S&P 500 ETF Trust (SPY).

Performance was assessed using four key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R²).

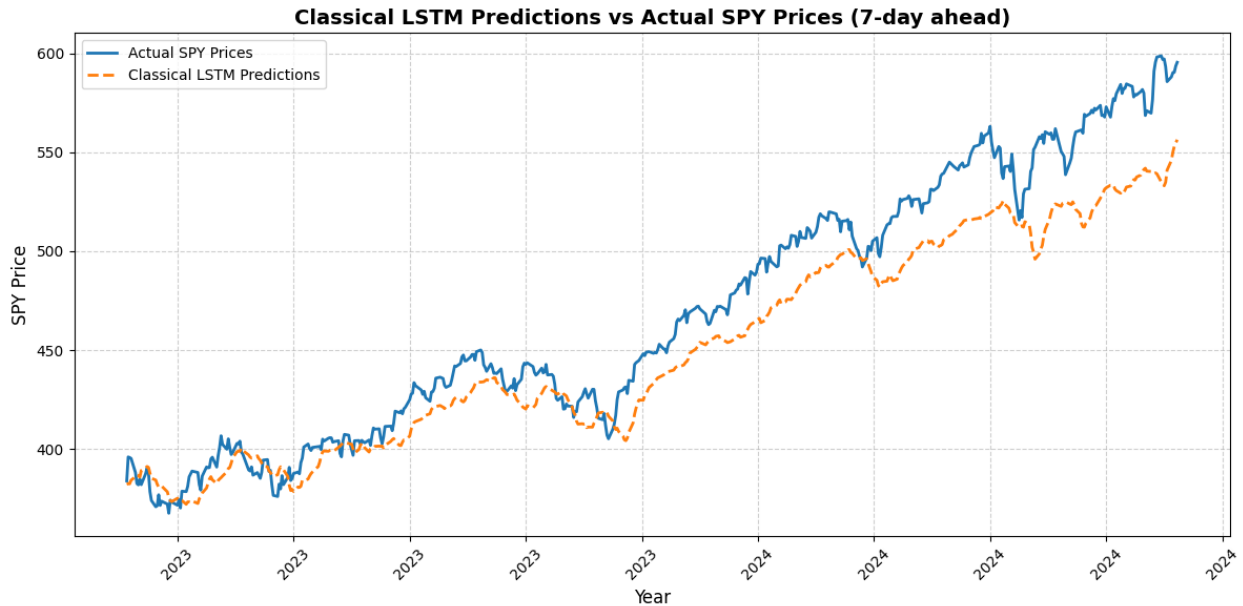


Figure 3: Classical LSTM Predictions vs SPY Prices

Classical LSTM Model Performance

The classical LSTM model demonstrated an ability to capture underlying trends in the SPY price data. The MAE of \$19.22 indicates that, on average, the model's predictions deviated from the actual SPY prices by approximately \$19.22. The RMSE of \$23.88, being more sensitive to larger errors, suggests that the model's performance was slightly impacted during periods of higher volatility. However, the relatively low MAPE of 3.83% and the high R^2 value of 0.86 underscore the model's overall strong predictive performance, particularly under stable market conditions. The classical LSTM model explained 86.51% of the variance in SPY prices, demonstrating its effectiveness in capturing temporal dependencies within the data, Figure 4.

Visual inspection of the predicted versus actual SPY prices (refer to Figure 3) reveals that the classical LSTM model generally aligned well with the actual price movements, particularly during periods of moderate volatility. However, the model exhibited some difficulty in accurately capturing sudden and extreme price fluctuations, contributing to the observed MAE and RMSE values.

The classical LSTM model's strengths lie in its simplicity and proven ability to model temporal dependencies in financial time series. It performed consistently well under most market conditions. However, its limitations were apparent in its struggle to adapt to abrupt market changes and outliers. The classical LSTM is well-suited for applications such as long-term trend

analysis and can provide actionable insights for portfolio rebalancing and long-term investment strategies. Furthermore, it serves as a valuable baseline for comparison with more complex models. While not ideal for high-frequency trading in volatile markets, its predictions can inform mid- to long-term algorithmic trading strategies.

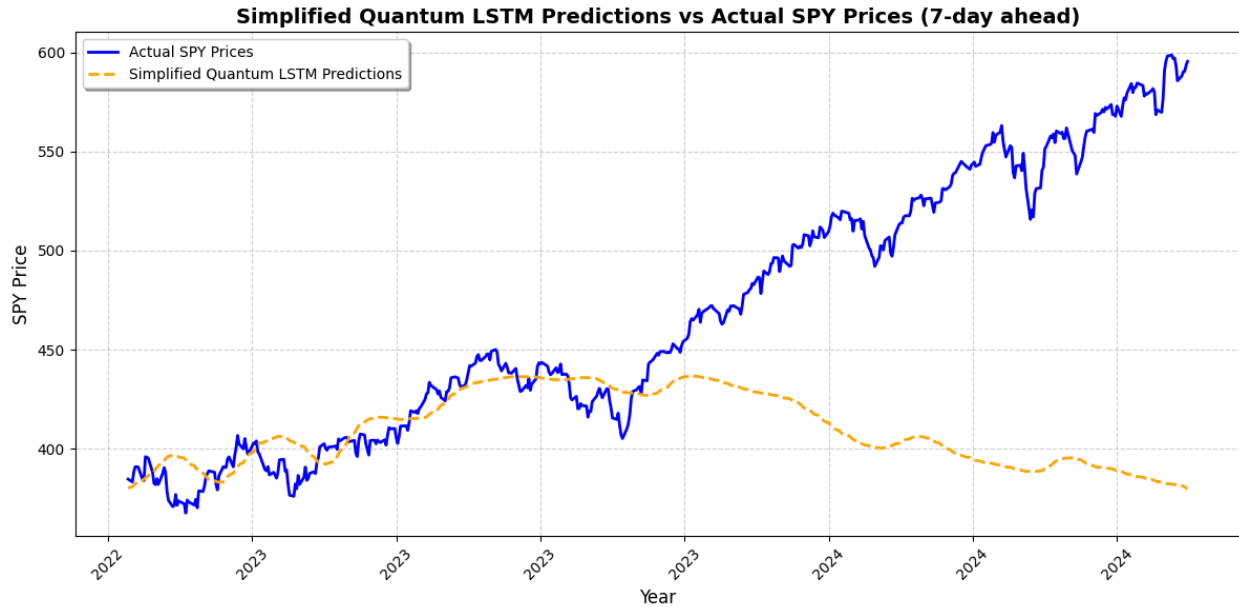


Figure 4: Simplified Quantum LSTM Predictions vs SPY

Simplified Quantum-Enhanced LSTM Model Performance

The simplified quantum-enhanced LSTM model exhibited performance limitations when applied to the SPY price prediction task. The significantly higher MAE of \$64.53 and RMSE of \$93.84 indicate that the model struggled to accurately predict SPY price movements, with large errors impacting its overall performance. The high MAPE of 12.21% further confirms its poor predictive accuracy relative to the classical LSTM and ensemble models. The negative R^2 value of -1.06 is particularly noteworthy, suggesting that the model performed worse than a naive model that simply predicts the mean of the target variable, Figure 5.

The simplified quantum-enhanced LSTM, while underperforming in its current form, serves as an important proof-of-concept for the integration of quantum circuits into financial modeling pipelines. Its primary weakness lies in the simplified quantum circuit design, which likely lacks the complexity to effectively capture the nuances of financial time series data (Chen, Yoo et al. 2022, Khan, Muzammil et al. 2024). Despite its current limitations, this model serves as a foundation for further exploration of quantum machine learning in finance. As quantum computing technology matures and more sophisticated quantum algorithms are developed, the

performance of such models is expected to improve. They may also have a place as a complement to other models.

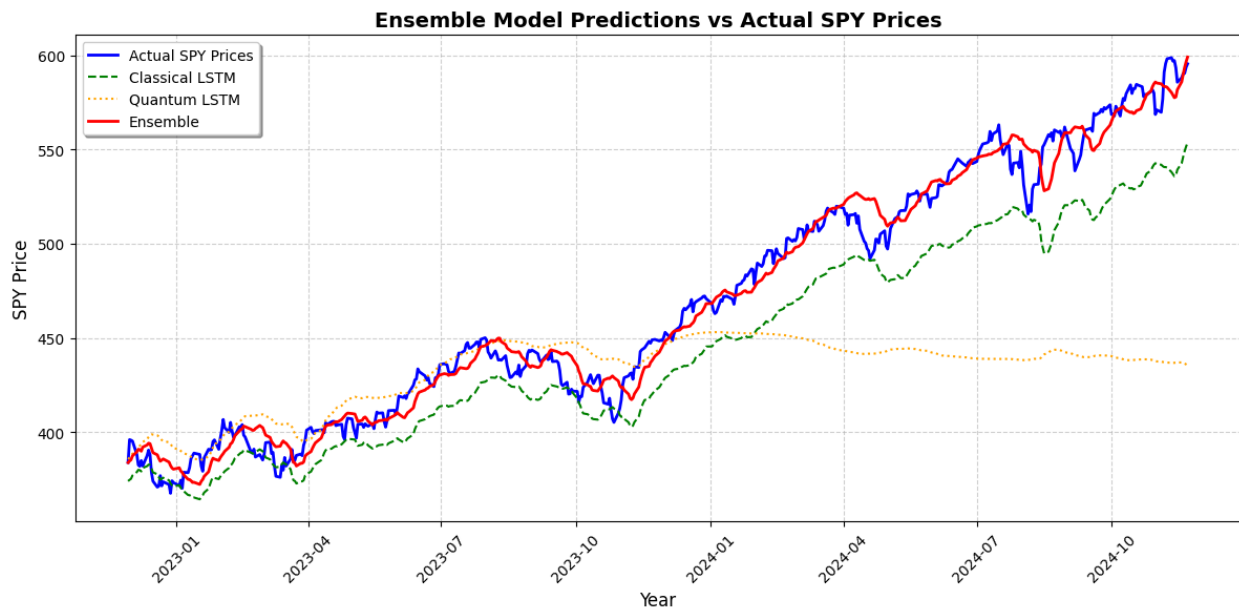


Figure 5: Ensemble Model Prediction vs Spy

Ensemble Model Performance

The ensemble model, combining the predictions of the classical LSTM and the quantum-enhanced LSTM, demonstrated superior performance compared to either of the individual models. The ensemble model achieved a remarkably low MAE of \$8.12 and RMSE of \$10.16, indicating a high degree of accuracy and consistency in its predictions. The MAPE of 1.75% further underscores its strong predictive power, with errors averaging less than 2%. The R^2 value of 0.97 is particularly impressive, demonstrating that the ensemble model explained 97.56% of the variance in SPY prices, outperforming both individual models by a significant margin.

Visual analysis of the ensemble model's predictions (refer to Figure 5) reveals a close alignment with the actual SPY price movements, even during periods of high volatility. The model effectively leveraged the strengths of both the classical LSTM, which provided a stable foundation during regular market conditions, and the quantum-enhanced LSTM, which potentially contributed unique insights into complex patterns. The ensemble model's strengths are in its superior accuracy. It combines the best of both worlds. However, it is a complex model and may be difficult to implement.

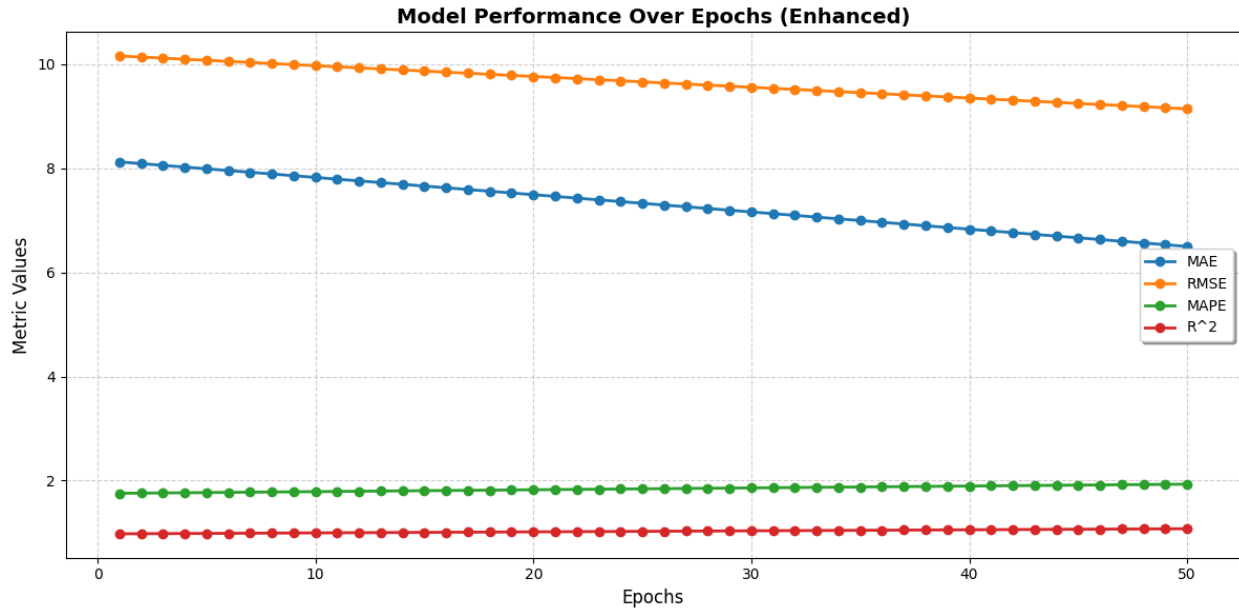


Figure 6: Model Drift Captured over Epochs

Data Drift Analysis and Implications

Analysis of input feature distributions using the Kolmogorov-Smirnov (KS) test revealed statistically significant drift across all features, with p-values below the established significance threshold ($\alpha = 0.05$). This finding indicates substantial changes in the underlying data distributions over time. Notably, features such as the VIX, SPY, and BTC adjusted closing prices exhibited extreme drift, suggesting a potential degradation in model performance if these shifts are not addressed. Similarly, macroeconomic indicators, including GDP Growth and Unemployment Change, also demonstrated significant drift, reflecting substantial shifts in the broader economic landscape. These findings underscore the necessity for periodic model retraining to incorporate emerging trends and maintain predictive accuracy, Figure 6.

To mitigate the impact of data drift, it is recommended to implement a continuous data drift monitoring system to regularly assess changes in feature distributions. In conjunction with monitoring, concept drift mitigation strategies should be adopted. These may include periodic retraining of the model using updated data to reflect the most current market conditions or employing adaptive learning techniques that allow the model to adjust to evolving data patterns dynamically.

As depicted in Figure 6., the ensemble model demonstrated consistent performance improvements over the training epochs. However, the substantial data drift observed in the input features highlights the critical need for ongoing monitoring and adaptive retraining to ensure the model remains accurate and reliable in the context of dynamic financial markets. The integration

of these monitoring and adaptation strategies is essential for sustaining the long-term effectiveness of advanced analytical models in financial forecasting.

Feature	KS Statistic	P-Value	Observation
SPY Adj Close	1.0000	0.0000e+00	Significant drift detected; requires attention.
SPY Log Returns	0.9996	0.0000e+00	Significant drift in return patterns.
VIX Adj Close	1.0000	0.0000e+00	Extreme drift due to changes in volatility index.
SDS Adj Close	1.0000	0.0000e+00	Consistent with SPY patterns; significant drift.
SDS Log Returns	0.9992	0.0000e+00	Indicates considerable return shifts.
BTC Adj Close	1.0000	0.0000e+00	Bitcoin's price dynamics show major drift.
BTC Log Returns	0.9996	0.0000e+00	Reflects shifting speculative behavior.
GDP Growth	0.8780	0.0000e+00	Detectable drift; possibly influenced by recessions.
Unemployment Change	0.9949	0.0000e+00	Strong drift, likely due to macroeconomic changes.
VIX_GDP_Interaction	0.9271	0.0000e+00	Interplay of volatility and GDP has shifted.

Figure 7: Model Drift Results

DISCUSSION

Financial forecasting has undergone significant evolution over the decades, from rudimentary statistical techniques to the incorporation of sophisticated machine learning (ML) and deep learning (DL) methods. Early approaches in the mid-20th century relied heavily on fundamental and technical analysis, using linear regression and time-series models like ARIMA to predict

market trends. While these methods offered some utility in stable market conditions, they often fell short in capturing the complexity and non-linearity of financial markets, particularly during periods of heightened volatility or structural change (Makridakis, Spiliotis et al. 2023).

The rise of computational power and data availability in the late 20th and early 21st centuries heralded the introduction of ML techniques, such as decision trees, random forests, and support vector machines (SVMs) (Ryll and Seidens 2019). These models marked a departure from traditional forecasting by emphasizing pattern recognition in large datasets. However, they remained constrained by their inability to model temporal dependencies effectively, an important consideration in financial time series data (von Krogh, Roberson et al. 2023).

Deep learning, particularly Long Short-Term Memory (LSTM) networks, emerged as a game-changer in the 2010s (Fischer and Krauss 2018, Sagheer and Kotb 2019). Designed to overcome the limitations of traditional recurrent neural networks (RNNs), LSTMs excel in modeling long-term dependencies, making them highly effective for time-series forecasting. This evolution aligns with a broader historiographical trend in financial forecasting: the shift from static, rule-based models to dynamic, data-driven approaches that leverage the complexity of financial markets (Fischer and Krauss 2018, Kulshrestha and Srivastava 2020).

For stock market predictions seven days out, these advantages are pronounced. In this study, the classical LSTM model achieved a Mean Absolute Error (MAE) of 19.22, Root Mean Squared Error (RMSE) of 23.88, Mean Absolute Percentage Error (MAPE) of 3.83%, and R-squared of 0.86. These results highlight the model's ability to capture short-term market dynamics (Oliver-Muncharaz 2020). The relatively low error rates and high R-squared value demonstrate the strength of LSTMs in identifying and extrapolating patterns from historical data.

The ensemble modeling approach adopted in this study represents a significant paradigm shift for the financial services industry. By combining the strengths of classical LSTM models with innovative quantum-inspired architectures, the ensemble model achieves a level of accuracy and robustness that surpasses standalone models (Rane, Choudhary et al. 2024).

The ensemble model's strong performance underscores its potential to transform investment analytics within RIA firms (Bhandari, Rimal et al. 2022, Majumder, Rahman et al. 2022). By bridging the gap between traditional methods and cutting-edge quantum-enhanced approaches, it offers a pathway for RIAs to adapt to the data-driven financial landscape.

This study also underscores the transformative potential of ensemble modeling in financial forecasting. By demonstrating the value of combining classical and quantum-inspired approaches, it provides a roadmap for RIAs seeking to modernize their investment analytics. The findings suggest that adopting such advanced techniques can improve forecasting accuracy, enhance portfolio management strategies, and ultimately deliver superior value to clients.

The significant performance improvements achieved by the ensemble model, particularly the reduction in MAE from 19.22 to 8.12 and the increase in R-squared from 0.86 to 0.97 highlight its paradigm-shifting potential. These results not only validate the efficacy of the ensemble approach but also challenge traditional forecasting paradigms, emphasizing the need for a broader industry-wide shift toward advanced analytics.

As the historiography of financial forecasting continues to evolve, this research positions ensemble modeling as a cornerstone of the next generation of financial services based analysis. By integrating advanced ML, DL, and quantum-inspired techniques, RIA firms can transition from reactive to proactive decision-making, achieving both competitive differentiation and enhanced client outcomes.

FUTURE RESEARCH CONSIDERATIONS

Moving forward, further research should focus on refining quantum circuit designs, exploring more sophisticated quantum algorithms, and integrating these techniques with other advanced machine learning approaches, such as deep reinforcement learning or other ensemble configurations. Investigating the application of these models to a wider range of asset classes, markets, and forecasting horizons will also be important considerations. Moreover, addressing the “black box” nature of these models through explainable AI (XAI) techniques will become increasingly important for fostering trust and adoption within the financial industry. By continuing to push the boundaries of research in this rapidly evolving field, we can unlock the full potential of advanced analytics to revolutionize financial forecasting and empower RIA firms to deliver superior value and service in an increasingly complex and data-driven world. Finally, the exploration into other open-sourced tools should be considered for further research as a means to lower costs while improving performance.

LIMITATIONS OF THE STUDY

This study provides insights into the application of advanced analytics for financial forecasting, specifically within the context of RIA firms. However, several limitations warrant consideration. Firstly, the quantum-enhanced LSTM model employed in this research utilized a simplified quantum circuit. This simplification was necessitated by the current limitations of available quantum hardware and the nascent stage of quantum algorithm development for complex financial applications. Consequently, the model may not fully capture the potential of quantum computing in finance, and future research employing more sophisticated quantum circuits and algorithms may yield different results.

Secondly, the scope of this study was primarily focused on the S&P 500 ETF (SPY) as a proxy for the U.S. equity market. While SPY serves as a broad market indicator, the findings may not be directly generalized to other asset classes, international markets, or individual securities exhibiting different risk and return characteristics. Moreover, the models were trained and tested

on historical data, and their performance might vary under different market regimes, especially during periods of unprecedented volatility. It is important to note that the specific time period used can influence the results and should be carefully considered when analyzing the model's performance. Additionally, the computational demands of training complex deep learning and quantum-inspired models are substantial. Although this study leveraged GPU acceleration, access to more powerful computing infrastructure could facilitate the exploration of more complex model architectures and larger datasets, potentially leading to improved performance.

Furthermore, while hyperparameter tuning was performed, the vast search space for optimal configurations suggests that the models may not be fully optimized. Further exploration of hyperparameter combinations, possibly using more advanced optimization techniques, could yield further performance enhancements. The study also utilized a specific set of features, including macroeconomic and technical indicators derived from price data. While this set was carefully selected, other potentially relevant features, such as sentiment data from news articles or social media, were not included. Incorporating such features could enhance model accuracy but would also introduce additional complexity in data collection and model development. Another limitation to this study is the models produced are still subject to the "black box" nature, meaning further research is needed to determine how these models arrive at their conclusions.

Finally, the study primarily focused on short-term (7-day ahead) forecasting. While relevant for many trading and investment strategies, these findings may not be directly applicable to longer-term forecasting horizons. Additionally, the analysis did not account for transaction costs, which can significantly impact the profitability of trading strategies based on model predictions. Incorporating realistic transaction cost models would provide a more accurate assessment of the practical applicability of the findings. Despite these limitations, this study contributes to the growing body of research on the application of advanced analytics in finance and provides a foundation for further investigation and development of more sophisticated and robust forecasting models. Future research should address these limitations, particularly by exploring more complex quantum circuits, expanding the dataset, incorporating additional features such as sentiment analysis, evaluating model performance across various market conditions and time horizons, and further exploring model interpretability.

CONCLUSION

This study investigated the application of advanced machine learning techniques, specifically ensemble learning and quantum-inspired methods, to enhance financial forecasting for Registered Investment Advisor (RIA) firms. The results provide evidence for the potential of ensemble learning, particularly when combined with nascent quantum-inspired techniques, to significantly improve the accuracy and robustness of financial forecasting models. The ensemble model, integrating a classical LSTM with a simplified quantum-enhanced LSTM, demonstrated

outsized performance across all evaluation metrics compared to the individual models. Its ability to achieve a low Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), coupled with a high R-squared value, underscores its effectiveness in capturing complex, non-linear market dynamics and providing reliable predictions, even during periods of high volatility.

These findings have significant implications for RIA firms seeking to leverage advanced analytics to refine their investment strategies, optimize portfolio allocation, and enhance risk management. The improved accuracy of the ensemble model can contribute to more informed decision-making, potentially leading to better investment returns and alpha generation. By providing more precise forecasts, these models can empower advisors to better navigate market fluctuations and tailor their recommendations to individual client needs and risk profiles.

While the simplified quantum-enhanced LSTM model did not outperform classical methods in this particular study, its inclusion represents an important step towards exploring the transformative potential of quantum computing in finance. The observed limitations of the quantum component highlight the current challenges associated with quantum hardware and algorithm development for complex financial applications. However, it also serves as a foundation for future research and highlights the potential for significant advancements as the field of quantum computation matures.

REFERENCES

- Ackerman, S., et al. (2020). "Detection of data drift and outliers affecting machine learning model performance over time." arXiv preprint arXiv:2012.09258.
- Ampatzoglou, A., et al. (2015). "The financial aspect of managing technical debt: A systematic literature review." Information and Software Technology **64**: 52-73.
- Asgarova, A. and F. Sattarova (2021). "Big Data Application in Finance." International Journal of Scientific Research and Management **9**(11): 91-98.
- Banker, R. D., et al. (2019). "Software Process and Business Value: An Empirical Study of Technical Debt and Firm Performance." MIS quarterly **43**(3): 897-924.
- Bayani, S. V., et al. (2024). "Interpretability and explainability in AI based financial models: A comprehensive review and future research directions." Expert Systems with Applications **242**: 122720.
- Bhandari, H. N., et al. (2022). "Predicting stock market index using LSTM." Machine Learning with Applications **9**: 100320.
- Brown, N., et al. (2010). Managing technical debt: Reducing friction in software development. 2010 ICSE Workshop on Managing Technical Debt.

- Bruder, B., et al. (2013). "Regularizing Mean-Variance Portfolio Allocation." Journal of Investing **22**(4): 66-76.
- Chen, S. Y.-C., et al. (2022). Quantum long short-term memory. ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE.
- Codabux, N., et al. (2017). "The financial aspect of managing technical debt: A systematic literature review." Journal of Systems and Software **132**: 192-219.
- Cohen, G. (2022). "AI and the Future of Finance: A Pragmatic View of Trends, Risks, and Opportunities." Journal of Financial Data Science **4**(2): 8-23.
- Conroy, N. (2012). Technical Debt as a Financial Instrument: A Case Study in the Adoption of Technical Debt. 2012 ICSE Workshop on Managing Technical Debt.
- Das, D., et al. (2021). Towards Automated Detection of Technical Debt in Machine Learning Systems. 2021 IEEE/ACM International Conference on Technical Debt (TechDebt).
- Dhingra, A. and A. Ashok (2021). Application of big data in the financial sector. Big Data Analytics: Methods and Applications: 59-82.
- Dudek, G. (2023). "Combining Forecasts of Time Series with Complex Seasonality Using LSTM-Based Meta-Learning." Engineering Proceedings **39**(1): 53.
- Durgapal, V. and S. Vimal (2021). Short Term Stock Price Prediction using Extreme Gradient Boosting. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).
- Fairley, R. E. and M. J. Willshire (2017). Technical Debt: A Risk Management Approach. 2017 IEEE/ACM International Conference on Technical Debt (TechDebt).
- Fang, J., et al. (2014). "Predictability of the simple technical trading rules: An out-of-sample test." Review of Financial Economics **23**(1): 30-45.
- Fischer, T. and C. Krauss (2018). "Deep learning with long short-term memory networks for financial market predictions." European journal of operational research **270**(2): 654-669.
- Frischbier, S., et al. (2019). The Data Challenge of the Financial Services Industry: Towards Data-Driven Innovation. Proceedings of the 27th European Conference on Information Systems (ECIS).
- Gökgöz, F. (2017). "CAPM and Three-Factor Model: A Comparison on the Istanbul Stock Exchange." Journal of Economics, Finance and Accounting **4**(2): 141-150.

- Gollapudi, S. (2015). Big Data in the Financial Services Industry: Challenges and Opportunities.
- Jothimani, D., et al. (2016). A hybrid approach for stock price prediction using empirical mode decomposition and artificial neural network. 2016 International Conference on Computation of Power, Energy Information and Communication (ICCPEIC).
- Khan, S. Z., et al. (2024). "Quantum long short-term memory (QLSTM) vs. classical LSTM in time series forecasting: a comparative study in solar power forecasting." Frontiers in Physics **12**: 1439180.
- Khurana, U., et al. (2018). Feature engineering for predictive modeling using reinforcement learning. Proceedings of the AAAI Conference on Artificial Intelligence.
- Kim, J., et al. (2021). "Robust Portfolio Optimization: An Overview." Journal of portfolio management **47**(5): 29-42.
- Kontopoulou, V. I., et al. (2023). "A Deep Learning Hybrid Framework for Time Series Forecasting: Combining Statistical, Deep Learning, and Meta-Learning Models." Expert Systems with Applications **213**: 118846.
- Kritzman, M. (2011). Markowitz's 'Portfolio Selection'. 50 Years of Modern Portfolio Theory, Emerald Group Publishing Limited.
- Kulshrestha, N. and V. K. Srivastava (2020). Synthesizing Technical Analysis, Fundamental Analysis & Artificial Intelligence—An Applied Approach to Portfolio Optimisation & Performance Analysis of Stock Prices in India. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), IEEE.
- Lai, T. L., et al. (2011). "Mean-variance portfolio optimization when means and covariances are unknown." The Annals of Applied Statistics **5**(2A): 798-823.
- Liu, J.-X., et al. (2023). "Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM." PeerJ Computer Science **9**: e1403.
- Ma, B., et al. (2021). Research on stock index prediction and portfolio optimization model based on machine learning. Journal of Physics: Conference Series.
- Majumder, A., et al. (2022). Stock market prediction: a time series analysis. Smart Systems: Innovations in Computing: Proceedings of SSIC 2021, Springer.

Makridakis, S., et al. (2023). "Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward." Journal of the Operational Research Society **74**(3): 840-859.

Mukhopadhyay, S. and S. Banerjee (2010). "Cooperating swarms: A paradigm for collective intelligence and its application in finance." International Journal of Computer Applications **6**(10): 31-41.

Munar, W., et al. (2014). Big Data Management in Financial Services: Results from a Delphi Study. Proceedings of the 22nd European Conference on Information Systems (ECIS).

Muskaan, P., et al. (2022). "Comparative analysis of GRU and Bi-LSTM for predicting the stock market." Procedia Computer Science **215**: 645-655.

Oliver-Muncharaz, J. (2020). "Comparing classic time series models and the LSTM recurrent neural network: An application to S&P 500 stocks." Finance, Markets and Valuation **6**(2): 137-148.

Orús, R., et al. (2018). "Quantum computing for finance: Overview and prospects." Reviews in Physics **4**: 100028.

Pasupulety, V., et al. (2019). Stock Price Prediction Using Technical Indicators and Sentiment Analysis. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC).

Ramasubbu, N. and C. F. Kemerer (2015). System Failures, Reliability, and the Cost of Technical Debt. 2015 48th Hawaii International Conference on System Sciences.

Rane, N., et al. (2024). "Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions." Studies in Medical and Health Sciences **1**(2): 18-41.

Rane, N. and N. Mehendale (2023). Convolutional Neural Network-Based Stock Trend Prediction. Proceedings of Third International Conference on Computing, Communications, and Cyber-Security (Lecture Notes in Networks and Systems, vol 577), Springer, Singapore.

Reddy, B. M. and B. S. Reddy (2023). "Big Data in Financial Services: A Review." International Journal of Advanced Computer Science and Applications **14**(1): 45-54.

Roberts, N., et al. (2014). Hidden technical debt in machine learning systems. Advances in Neural Information Processing Systems, 27.

- Rundo, F., et al. (2019). "Machine learning for quantitative finance applications: A survey." Applied Sciences **9**(24): 5574.
- Ryll, L. and S. Seidens (2019). "Evaluating the performance of machine learning algorithms in financial market forecasting: A comprehensive survey." arXiv preprint arXiv:1906.07786.
- Sagheer, A. and M. Kotb (2019). "Time series forecasting of petroleum production using deep LSTM recurrent networks." Neurocomputing **323**: 203-213.
- Selvin, S., et al. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 international conference on advances in computing, communications and informatics (icacci), IEEE.
- Siame-Namini, S. and A. Namin (2018). Forecasting Economics and Financial Time Series: ARIMA vs. LSTM.
- Siame-Namini, S., et al. (2018). A comparison of ARIMA and LSTM in forecasting time series. 2018 17th IEEE international conference on machine learning and applications (ICMLA), IEEE.
- Siaw, R., et al. (2014). Portfolio optimization using ARIMA and GARCH models. 2014 IEEE Conference on Systems, Process & Control (ICSPC 2014).
- Sridevi, S., et al. (2011). Imputation for the analysis of missing values and prediction of time series data. 2011 international conference on recent trends in information Technology (ICRTIT), IEEE.
- Sugadev, K., et al. (2023). Portfolio Optimization and Recommendation: A Deep Learning Approach. 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT).
- Théate, T. and D. Ernst (2021). "An application of deep reinforcement learning to algorithmic trading." Expert Systems with Applications **173**: 114632.
- Viswanathan, K. and B. A. Stephen (2020). Ensemble model for short term forecasting of NIFTY 50 using closing value. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS).
- von Krogh, G., et al. (2023). Recognizing and utilizing novel research opportunities with artificial intelligence, Academy of Management Briarcliff Manor, NY. **66**: 367-373.
- Warin, X. (2021). "Deep learning for covariance matrices." Risks **9**(3): 58.

Yang, W. (2023). Mean-Variance Portfolio Optimization Based on ARIMA and GARCH Models.
2023 7th International Conference on E-commerce, E-Business and E-Government.

Tech For Health: Understanding and Identifying Quality Health Information Online

Loreen Powell, Ph.D., Marywood University, Scranton

Gwendolyn Powell, Penn State University, University Park

Chaza Abdul, Marywood University, Scranton

Ronda Mariani, Commonwealth University of PA, Bloomsburg

Abstract

The recent health pandemic and the rapid evolution of digital technologies has led to an increase in the amount of poor-quality, biased, and misleading information available online. As a result, many seniors learned to use digital technologies to communicate with their healthcare providers and learn health information. However, in 2022 the Health on the Net (HON) Foundation, which previously provided certification for trustworthy health websites, was dismantled. Thus, leaving everyone without a certification to look when searching for quality healthy websites. This paper proposes a five-step framework, based on the HON Code and the National Institute on Aging guidelines, to help seniors identify reliable health information online.

Keywords: *HON Code, quality digital content, health information, seniors*

Brief Review of Literature

As technologies continue to evolve at exponential rates, so is the amount of poor-quality, biased, and misinformation found online. Furthermore, the recent health pandemic helped to further push American society and the healthcare sector to embrace digital technologies. As a result of the increase in digital technologies used in healthcare to communicate with doctors and nurses, many elderly felt forced to learn how to use digital technologies and go online (Yeung, et al.,2023). In an effort to help many elderly learn digital technologies and how to search the web, many non-profit organizations, like Partners for a Healthy Community, Elysburg, started digital technology education programs (PHC, 2024). While these programs were effective in getting the elderly online, they may not have addressed many valuable insights on how to recognize quality health information online without the Health on Net (HON) Code certification. Furthermore,

if they addressed how to find quality health information online via the HON Code certification it would look different today as the Health on the Net (HON) Code certification no longer exists.

From 1996 to 2022, the Health on the Net (HON) Foundation was a reliable organization aimed to ensure reliable online health content. They provided a certification to educate website publishers on the ethical principles like the HON Code to health and medical information disseminated online. They also helped monitor and provide health guidelines to identify quality health information online (International Medical Informatics Association; 2025).

Today, the HON Foundation no longer exists due to funding challenges and the growing complexity of the digital health landscape. Thus, from 2022 to present, there has not been an official organization encouraging, focusing, and certifying high quality health information online (Boyer et al., 2018). The HON Code certificate that once was an indicator for seniors to recognize quality health information no longer can be found. As a result, there is a need to retrain seniors on how to understand how to find quality health information online. This need is also recognized by the National Institute on Aging (NIA). As a result, they also provide some tips for seniors (National Institute on Aging, 2023). However, their website, which provides tips on understanding quality websites, is very lengthy and dense with text. The goal of this paper is to provide a short and concise framework, based upon the previous HON Code and the NIA quality website (<https://www.nia.nih.gov/health/healthy-aging/how-find-reliable-health-information-online>), for non-profit organizations and seniors to understand how to identify quality health information online. The remainder of this paper provides a method, proposed framework, conclusion, limitations, and future research.

Method

The proposed Framework is adapted from the previous HON Code and the National Institute on Aging's Quality Resource website (<https://www.nia.nih.gov/health/healthy-aging/how-find-reliable-health->

[information-online](#)), to help seniors with understanding what to look for regarding quality health information online. The Delphi approach was used to create the proposed framework. Two of the authors were identified as experts within the healthcare informatics field due to healthcare publications and higher education teaching experience. Leveraging the author's extensive experience and the existing body of literature, they collaboratively refined the framework through an iterative process of feedback and revisions until they reached consensus. Next, the framework was shared with all authors, and it underwent additional reviews and iterations until consensus was reached by all authors. This methodical approach ensured a robust and well-founded framework, reflecting their expertise.

Proposed Framework

The proposed framework was designed to be a simple five step framework for aiding seniors in understanding health information online. Listed below in figure 1 is the proposed framework. It is important to note that each step contains a specific statement as to what seniors may want to consider. This statement is in black text. Next, a simple suggestion on how to look for the context of each statement is listed in blue text.

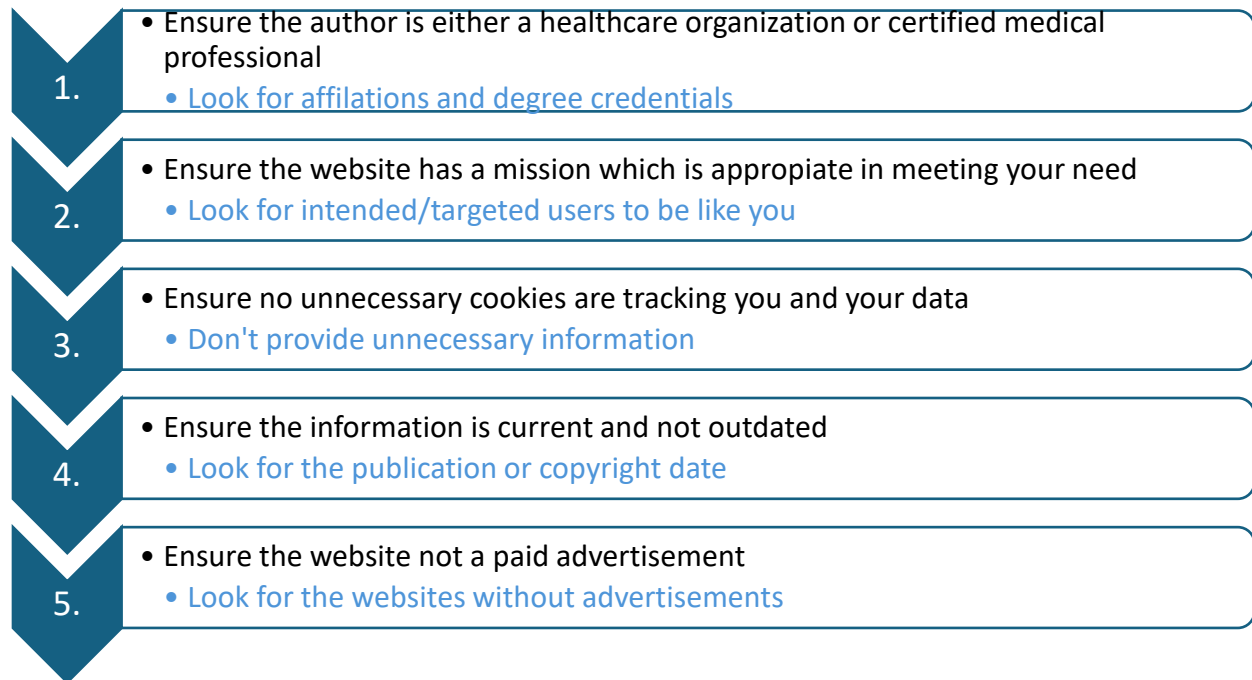


Figure 1: Proposed Framework

Conclusion, Limitations, & Future Research

This paper provided the first step in forming a simple framework for seniors to understand and identify quality health information online. It is important to note that this paper is not without limitations. First, the study is limited to a framework geared towards seniors. Secondly, the framework was created based upon the HON Code, the NIA quality website and author's experience. Additional validity testing should also be added. Finally, the authors recommend future research on addressing the limitations and perhaps applying the framework to various generations in addition to the elderly. Regardless, this paper provides practical implications for non-profit organizations looking to train or retrain seniors on using digital technologies and content. The paper also has practical implications for seniors and healthcare organizations.

References

Boyer, C, Selby, M, Scherrer, JR, & Appel, RD. (2018) The health on the Net Code of Conduct for medical and health websites. *Computer Biol Med.*;28(5):603-10. Doi: 10.1016/s0010-4825(98)00037-7. PMID: 9861515.

International Medical Informatics Association (2025). Health On the Net Foundation (HON). <https://imia-medinfo.org/wp/health-on-the-net-foundation-hon/>

National Institute on Aging (2023). How To Find Reliable Health Information Online. <https://www.nia.nih.gov/health/healthy-aging/how-find-reliable-health-information-online>

Partners for a Healthy Community (2024). Facebook Page for Partner for a Healthy Community -Elysburg PA. <https://www.facebook.com/people/Partners-For-a-Healthy-Community/100077659871261/>

Yeung, AWK, Torkamani, A, Butte, AJ, Glicksberg, BS, Schuller, B, Rodriguez, B, Ting, DSW, Bates, D, Schaden, E, Peng, H, Willschke, H, van der Laak, J, Car, J, Rahim, i K, Celi ,LA, Banach, M, Kletecka-Pulker, M, Kimberger, O, Eils, R, Islam, SMS, Wong, ST, Wong, TY, Gao, W, Brunak, S, & Atanasov, AG. (2023). The promise of digital healthcare technologies. *Front Public Health*;11:1196596. Doi: 10.3389/fpubh.2023.1196596. PMID: 37822534; PMCID: PMC10562722.

Real Time Analytics System Design: An Integration of AWS Kinesis, Apache Pinot, and Apache Superset

Zhengzhong Shi

University of Massachusetts at Dartmouth
285 Old Westport Road, Dartmouth, MA 02747, USA
zshi@umassd.edu

Abstract

With the critical role played by the real time analytics system in firms, MBA students need to have a better understanding of how the real time analytics system works and its architecture. This paper intends to design a real time analytics system with cloud computing technologies to demonstrate the different components in the system and how they work together. Through working on this demo system, MBA students can develop a deeper understanding of the real time analytics system and gain valuable hands-on experiences. Hopefully, this experience may better position them in the competitive job market.

Keywords: Real Time Analytics, Architecture, Apache Pinot, Apache Superset, AWS Kinesis, Cloud Computing, MBA.

Introduction

With the advancement of mobile and cloud computing technologies, real-time analytics is increasingly playing a critical role in empowering companies manage their internal operations and customer services. “For a company like Uber, real-time data is the lifeblood of both their customer-facing and internal services” [1]. On the one hand, restaurant managers who are customers of UberEATS can not only check the sales of past seven days but also the sales of the current day itemized based on their menus to more effectively manage inventory and employee schedules [3]. On the other hand, APP developers inside Uber can take advantage of the real-time app crash analytics to reduce the mean time to crash detection and fix bugs in a timely manner [2]. Further, Stripe, as the payment processing infrastructure operator, is dependent on Apache Pinot to produce real-time analytics supporting both user-facing applications and internal use cases, processing more than 250 million API requests per day, with a peak of 13,000 requests per second [4]. This paper is to design a real time analytics system with multiple cloud-based components to demonstrate to MBA students how real time analytics system works in order to improve their understanding of the role of data analytics in business and better position them in the competitive job market.

The Solution Architecture

For a real-time analytics system, it has multiple components as indicated in Figure 1. Event producers are such as various mobile apps (such as Uber EATs, Uber EATs Orders, Uber -Request a Ride, Uber – Driver, and Amazon) constantly producing data. Streaming platforms are such as AWS Kinesis and Apache Kafaka responsible for delivering mobile data to the real time database such as Apache Pinot and Druid. Frontends can be the same apps as those producing data and special visualization systems (such as Apache Superset and Tableau). In the system we are going to design, services such as AWS Kinesis, Apache Pinot, Apache Superset are chosen. In addition, a python program is used to simulate a data source.

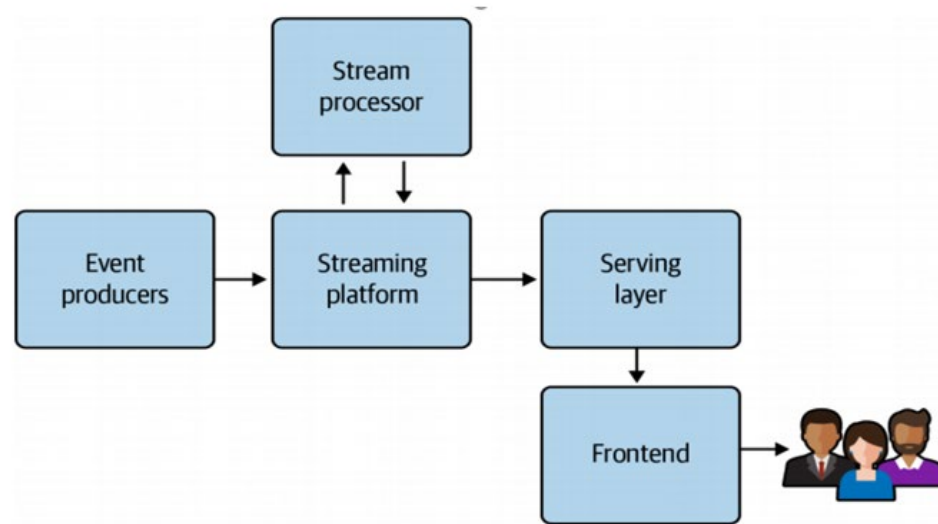


Figure 1: Components of a Real-Time Analytics System (Adopted from [5])

AWS Kinesis is a fully managed service responsible for ingesting, buffering, and processing streaming data in real time to derive insights with low latency [6]. Apache Pinot is a real time database system “for lightning-fast insights, effortless scaling, and cost-effective data-driven decisions” [7]. It has multiple components including such as brokers, controllers, servers, deep storage, zookeeper, and minion workers (Figure 2). Servers are responsible for storing data in segments with the columnar format and they support various indexing mechanisms such as the inverted index and the star-tree index. Brokers are the interface between SQL users and the database, and they consult controllers to get to know which servers should be used to execute the received queries. A controller plays the role of cluster manager responsible for node recovery, segment assignment, and schema management. Zookeeper [10] represents a cluster of computers responsible for electing the active controller and storing data for the controller. Minion workers are used for computing intensive jobs such as segment creation, merging, and purging.

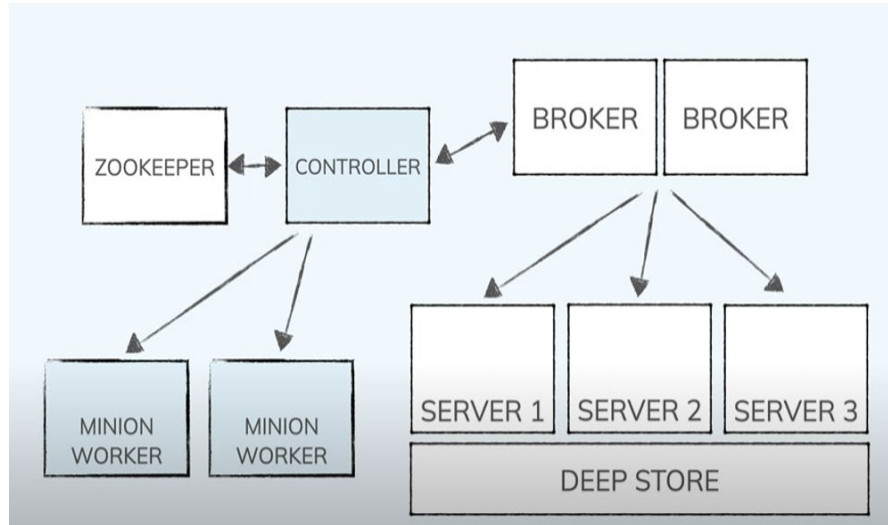


Figure 2: Apache Pinot Architecture (Adopted from [5])

Apache Superset is “an open-source modern data exploration and visualization platform” [8]. It integrates with a variety of data sources such as Apache Pinot, AWS Athena, AWS DynamoDB, Apache Spark SQL, Google Sheets, MySQL, SAP Hana, etc.

Key Configuration Procedures

Step 1: Installing and Configuring the Superset Server. An EC2 instance (in AWS) is used to set up the superset server following Superset documentation with necessary customization. The type of the instance should be at least t2.large if the sample dashboards in Superset are to be played with. Detailed procedures could be found at [11]. A virtual environment needs to be created to set up the python working environment. Configuration procedures can be found at [12]. Environmental variables such as `SUPERSET_CONFIG_PATH` and `FLASK_APP` need to be set up and the `superset_config.py` should be created with the `SECRET_KEY` configured. The connector for Apache Pinot should also be installed [13].

Step 2: Setting up the AWS Kinesis Stream and IAM role. A console command can be used in the cloud shell of AWS to create a Kinesis Stream [9]. Besides a Kinesis Stream, an IAM role allowing Apache Pinot to retrieve data from this stream should also be created and security credentials are used to set up the Apache Pinot database system.

Step 3: Setting up the Apache Pinot using the Data Manager (Figure 3) from startree.ai. The Data Manager from startree.ai has a free tier which is good for students to practice using Apache Pinot.

The data manager has a guided workflow (Figure 4) to set up both the connection with Kinesis and the data extraction from the stream. In particular, json strings could be easily processed by using customized functions (Figure 5). Sample data is generated by using a python program (application-metrics-generator.py) from [9].

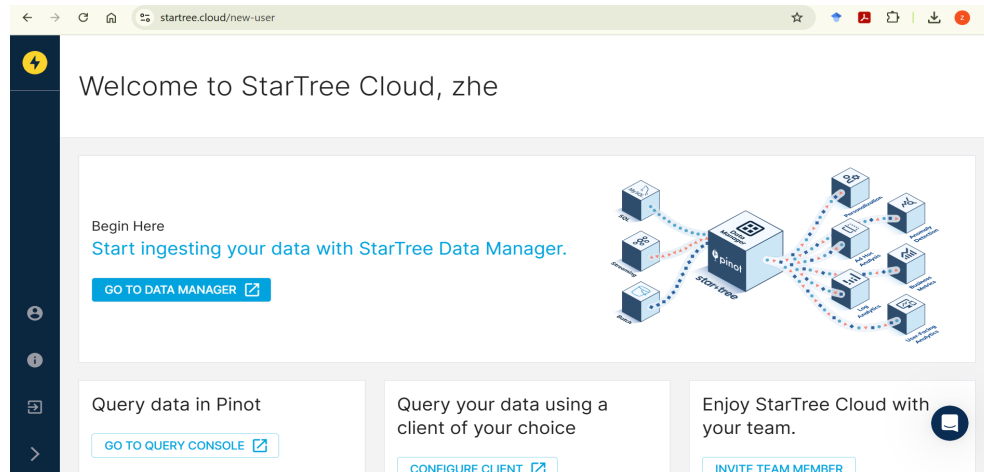


Figure 3: Data Manager from Startree.ai

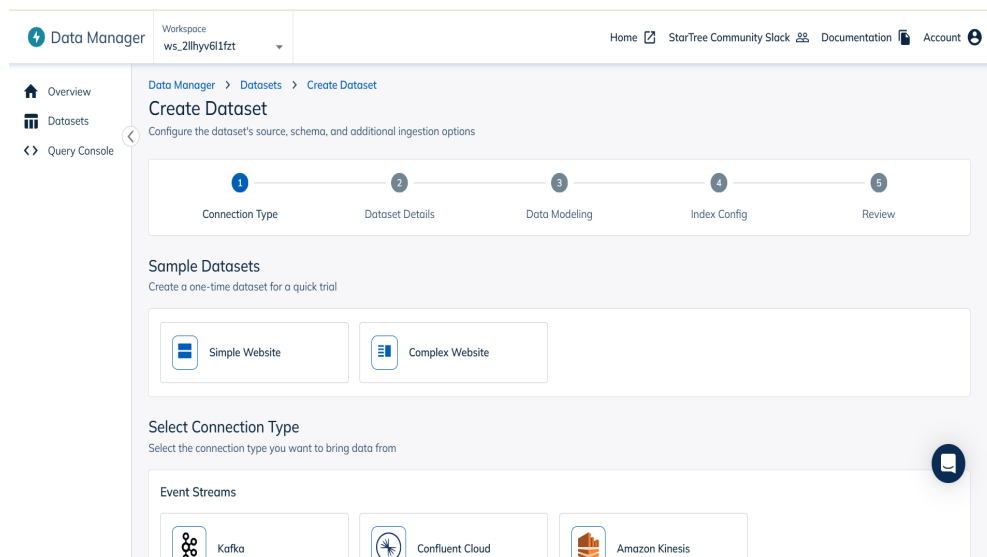


Figure 4: Data Manager Set Up Flow

After a successful setup, the data manager generates a SQLAlchemy URL which is used in setting up Superset to enable communications between the Superset and the Apache Pinot database (as managed by the startree.ai data manager).

Step 4: This step is to insert the SQLAlchemy URL from the previous step into Superset so that it can communicate with Apache Pinot (managed by the startree Data Manager). In addition, the chart, the dashboard, and the auto refresh of the dashboard should be configured. Figure 6 shows the dynamic chart with data coming the python program (which goes through the AWS Kinesis stream and the Apache Pinot Database).

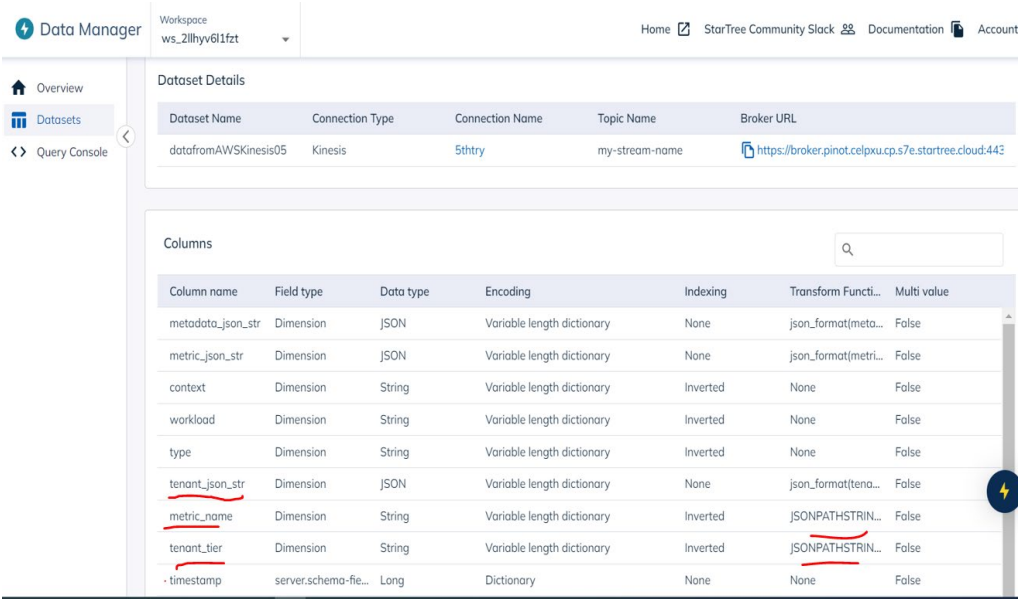


Figure 5: Kinesis Stream Data Extraction Mapping in The Data Manager

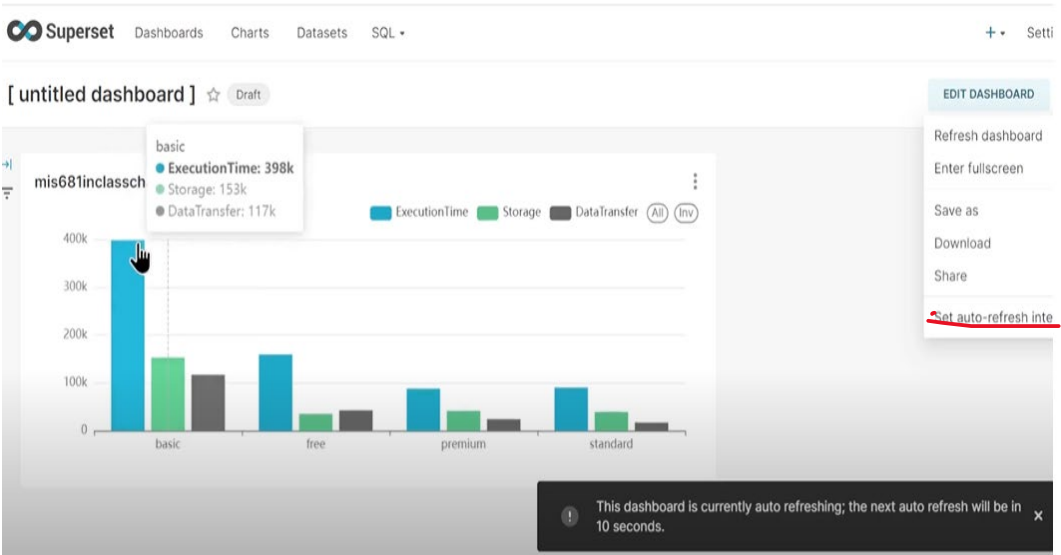


Figure 6: Dynamic Chart in Superset (Data Coming from Apache Pinot)

Conclusion

This paper discusses the architecture of a real time analytics system and the procedures to apply various cloud services such as the AWS Kinesis Stream, the AWS IAM role, the startree's Data Manager (for Apache Pinot), and the Apache Superset to set up a demo system. Through this hands-on experience, MBA students may better understand the architecture of a real time analytics system in the cloud computing platform, making them feel more confident in today's competitive job market. Future research could be to survey students to understand the effect of this demo system.

References

1. "Uber: Serving Real-Time App Crash Analytics While Saving \$2M+ With Apache Pinot", <https://startree.ai/user-stories/uber-serving-real-time-app-crash-analytics-while-saving-2m-with-apache-pinot>, Accessed on December 21, 2024.
2. "Real-Time Analytics for Mobile App Crashes using Apache Pinot", <https://www.uber.com/en-CO/blog/real-time-analytics-for-mobile-app-crashes/>, Accessed on December 21, 2024.
3. "Engineering Restaurant Manager, our UberEATS Analytics Dashboard", <https://www.uber.com/blog/restaurant-manager/>, Accessed on December 21, 2024.
4. "Stripe's Journey to \$18.6B of Transactions During Black Friday-Cyber Monday with Apache Pinot", <https://startree.ai/user-stories/stripe-journey-to-18-b-of-transactions-with-apache-pinot>, Accessed on December 21, 2024.
5. Mark Needham, "Building Real-Time Analytics Systems: From Events to Insights with Apache Kafka and Apache Pinot" Publisher : O'Reilly Media; 1st edition (October 24, 2023), ISBN-10 : 1098138791, ISBN-13: 978-1098138790.
6. AWS Kinesis, <https://aws.amazon.com/kinesis/>, Accessed on December 21, 2024.
7. Apache Pinot, <https://pinot.apache.org/>, Accessed on December 21, 2024.
8. Apache Superset, <https://superset.apache.org/>, Accessed on December 21, 2024.
9. AWS GitHub Repository, <https://github.com/wahab-io/apache-pinot-on-aws>, Accessed on December 21, 2024.
10. ZooKeeper, <https://zookeeper.apache.org/>, Accessed on December 21, 2024.
11. <https://superset.apache.org/docs/installation/pypi>, Installing Superset from PyPI, Accessed on December 21, 2024.
12. <https://superset.apache.org/docs/configuration/configuring-superset/>, Configuring Superset, December 21, 2024.
13. <https://superset.apache.org/docs/configuration/databases>, Connecting to Databases, December 21, 2024.

Managing Product Innovation: Lessons from Aviation History

Ram Gopalan

School of Business

Rutgers, The State University of New Jersey

Camden, NJ 08102

Email: ram.gopalan@rutgers.edu

Cite as:

Gopalan, R. (2025), Managing Product Innovation: Lessons from Aviation History, *Proceedings of the NEDSI (Northeast Decision Sciences Institute) conference, Hershey, PA, March 2025*.

Abstract

The study of business history can offer important lessons for managers, providing ideas for managing crises and also providing insights for managing complex product development efforts. The field of aviation in particular has accumulated a rich and colorful history – this paper reviews some critical incidents from aviation history, using them as an engine to generate a useful set of lessons and a checklist for innovation managers.

Keywords: Management of innovation, new products management, aviation history.

The product innovator's dilemma: to make a big leap or take small steps?

“We choose to go to the Moon in this decade and do the other things, not because they are easy, but because they are hard; because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one we intend to win.”—President John F. Kennedy.

Those lines were spoken by President Kennedy at an address made at Rice university on September 12th, 1962. As a consequence of that speech, Neil Armstrong walked on the moon on July 20, 1969, some seven years later. However at the time of his speech even President Kennedy could not have anticipated the moon landing and taken it for a certainty. In the very same speech, the President acknowledges that the project would require “a rocket made of new metal alloys, some of which have not yet been invented, capable of standing heat and stresses several times more than have ever been experienced (Kennedy (1962)).”¹ The historic speech adumbrated several other ground truths about high-technology innovation: space exploration could do good or go sideways, spin out of control and do ill. The project would require enormous commitments of money and resources, but revitalize the scientific and engineering communities in the United States, particularly around Houston. Salaries and expenses were

¹ Kennedy, J.F. (1962). historic speech at Rice university, 12th September 1962, **Digital identifier:** [USG-15-29-2, https://www.jfklibrary.org/learn/about-jfk/historic-speeches/address-at-rice-university-on-the-nations-space-effort](https://www.jfklibrary.org/learn/about-jfk/historic-speeches/address-at-rice-university-on-the-nations-space-effort), downloaded January 26, 2024.

estimated to be about \$60 million annually (Kennedy (1962)) and over \$200 million would be invested in plant and laboratory facilities. (see 1) In a nutshell, the President wanted to back a project fraught with risk but that would represent a big leap for the United States when compared to competitors like the Soviet Union. Moreover, success was not guaranteed. The lead times for even measuring progress and success were very high; he may not have known till the end of the decade whether it was actually going to work. The timing of the speech was fortuitous: the President had already completed a year and a half in office and maybe had another two years before he sought reelection. Since results were only guaranteed by the end of the decade, a failed moonshot may not have adversely affected his career (alas *kismet* intervened in 1963, making this discussion moot). Nevertheless, President Kennedy made the *choice* to back the moonshot. Most innovation managers face a similar situation – to bet the store on a single gigantic project, or seek small wins; and they must also decide whether to go out on the limb and promise delivery of results and publicize a deadline with lots of fanfare. Shah (2023) strongly advocates for ambitious “big bets” that can improve people’s lives and provides a disciplined process for managing such bets.² When Mr. Shah was in charge of USAid, he managed several crises, including the Haiti earthquake. He recommends “being optimistic, thinking bigger, finding fresh, innovative solutions, recruiting and retaining unlikely partners and fiercely measuring results, learning, and persisting”.

² Shah, Rajiv (2023). *Big Bets: How Large-Scale Change Really Happens*, Simon Element, October 2023. The quote was downloaded from <https://www.rockefellerfoundation.org/big-bets/book/> on February 5th, 2024.

Product innovation does not always have to be expensive

Now let us consider another important day in the history of aviation – December 17th, 1903. On this day, the day the first powered flight took place and it lasted only twelve seconds and traversed a mere 120 feet. The lucky pilot was Orville Wright. A few days earlier, his brother Wilbur had actually won the coin toss that would decide the identity of the history-making pilot at the helm of the Wright Flyer (McCullough (2016)), but as luck would have it, poor Wilbur crashed immediately upon take off.³ In flying as well as in product management, you might be able to take off easily and start a project, but it requires a bit of luck to land safely. In the very next attempt, on December 17th, Orville made history with his turn. In the short span between 1900 – 1930, aviation evolved at an exponential rate, comparable to the spread of AI and machine learning today. Very similar to the passion and leadership exuded by President Kennedy, the Wright brothers were the principal driving force behind powered flight. There was one important difference however when compared to space exploration – the Wright brothers were freelancing entrepreneurs and sought very little government backing. Indeed, powered flight which is perhaps a bigger and more ambitious a goal than even space exploration, cost very little to develop. Terwiesch and Ulrich (2009) also point out that a company's R&D spending may not be correlated with sales growth.⁴

Having a plan to fund the burn rate for product development

³ McCullough, David. (2016). The Wright brothers. Reprint edition ISBN: 978-1476728759.

⁴ Terwiesch, C and Ulrich, K.T. (2009). Innovation Tournaments, Harvard Business Press.

The Wright brothers started a bicycle shop in Dayton and this venture took off, enabling them to place Charlie Taylor (the same man who designed the engine for the Flyer) in charge of the business while they experimented with flying. Similarly, at the time Lindbergh considered the Atlantic crossing, he made money as an Air Mail pilot as well as by barnstorming (i.e., flying stunts for a crowd – Berg(1998)).⁵ Lindbergh’s plane called “The Spirit of St. Louis” had a set of backers from the same city and the team also had an agreement with the New York Times – Lindbergh would provide exclusive interviews to the paper if he survived the Atlantic crossing (Berg (1998)).(see 5) All of these factors helped to fund his adventure. In 1903, the Wright brothers had a fierce rivalry with Samuel Langley of the Smithsonian institution in their race to conquer the air. Mr. Langley’s efforts were well funded by the government, while the Wright brothers bootstrapped their efforts and spent far less money. Wilbur turned down offers to invest in the Wright Flyer (McCulloch (2016)).(see 3) Despite this imbalance in the level of resources, the brothers succeeded in December 1903, while Langley gave up after two failed attempts on October 7th and December 8th of the same year. Luckily, product success does not seem proportional to the amount of resources invested. The same airmail program that helped Lindbergh pay the bills initially was later investigated by the government when allegations of corruption emerged as to the granting of airmail contracts. Mungiu-Pippidi (2015) more recently also warns of the harmful effects of government corruption on innovation as a whole.⁶

⁵ Berg, A.S. (1998). Lindbergh. Putnam books. ISBN: 978-0399144493.

⁶ Mungiu-Pippidi, A. (2015). Good governance powers innovation, *Nature*, Vol. 518, February 2015.

Let the experiment be made!

Innovation managers need to constantly test various combinations of product features. Often, product performance under stressed conditions can be resolved with a simple, direct experiment that cuts to the heart of the matter. Benjamin Franklin who is credited with the invention of the lightning rod (among many other extremely useful inventions) popularized the phrase “Let the experiment be made”. He described an experiment to prove the efficacy of lightning rods in letters to his agent Collinson in London (this happened about 1750). On May 10, 1752, the Comte de Buffon and Thomas-Francois D’Alibard constructed a sentry box in the village of Marly on the outskirts of Paris and were able to draw electrical sparks from a storm cloud (Isaacson (2004)).⁷ Franklin’s own famous kite experiment followed shortly thereafter, in June of 1752. Franklin enlisted his son William to help fly the kite – Franklin was 46 years old and William was 21, so age may have nothing to do with innovation prowess.

In more recent times, on January 28, 1986, the Challenger space shuttle went up in flames soon after takeoff. The incident weakened confidence in NASA and the Rogers commission was put together to figure out the root cause of the explosion. After much scrutiny, it was thought that a product called an O-ring was not performing up to design standards in cold weather. How to prove this? The colorful Nobel prize winning physicist Richard Feynman came to the rescue

⁷ Isaacson, Walter (2004). Benjamin Franklin: An American Life, Simon and Schuster paperbacks, ISBN-13: 978-0-684-80761-4.

(McDonald and Hansen (2009)).⁸ On live TV, he dipped the O-ring into a glass of ice water and demonstrated that the product did not function as expected in cold weather.

The experiments of Franklin and Feynman had enormous impact, but neither was expensive to pull off. Companies that do most of their business via e-commerce, such as Amazon and AirBnB need to constantly experiment to optimize price levels and product features. Kohavi et al. (2020) discuss toolkits for experimentation in an online setting in detail.⁹ Finally, product innovators are curious enough to perform the experiment in the first place and they generally have a wide range of interests: Franklin, besides his lightning rod also invented the catheter and the Franklin stove; the Comte de Buffon is also famous for creating the “Buffon needle problem” in probability theory and he was also a well known naturalist with strong evolutionary ideas. This trait of being open to ideas from many fields may have been helpful to the early inventors and in recent times, even AI and machine learning have been enlisted to provide ideas for exploration (Wang et al. (2023)).¹⁰ Banerjee and Duflo (2012) have also championed the use of

⁸ McDonald, A.J and Hansen, J.R. (2009). Truth, Lies and O-Rings: Inside the space shuttle *Challenger* disaster, University of Florida press, ISBN-13: 978-0-8130-3326-6.

⁹ Kohavi, R, Tang, D and Xu, Y. (2020). Trustworthy online controlled experiments: A practical guide to A/B testing, Cambridge University Press, ISBN-13 : 978-1108724265.

¹⁰ Wang, H, Fu, T, Du, Y, Gao, W, Huang, K, Liu, Z, Chandak, P, Liu, S, Van Katwyk, P, Deac, A, Anandkumar, A, 13, Bergen, K, Gomes, C.P., Ho, S, Kohli, P, Lasenby, J, Leskovec, J, Liu, T.Y., Manrai, A, Marks, D, Ramsundar, B, Song, L, Sun, J, Tang, J, Veličković, Welling, M,

experiments and randomized control trials as a mechanism to develop policy levers to fight poverty, an effort that led to their being awarded the Nobel prize along with Michael Kremer in 2019.¹¹

Creating the product team: Lone wolf innovation or collaboratives of experts?

Often, we associate a single, colorful, dynamic and passionate individual with an invention or a company. Edison calls up the electric bulb, and Steve Jobs and Bill Gates are mastheads for Microsoft and Apple. Certainly, product innovation can benefit from the passionate vision and energy of a few key individuals. However not all breakthrough products result from a lone wolf's effort. Nielsen (2012) comments on the era of networked science and provides several specific examples of progress resulting from the efforts of *collections* of individuals.¹² Cornell university's laboratory for ornithology has launched project e-bird, allowing amateur bird watchers to upload their observations of a specific species, while tagging time and location. Such distributed data collection provides an invaluable repository of information about birds and

Zhang, L, Coley, Bengio, Y, Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence, *Nature*, August, 2023.

¹¹ Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty, Public Affairs, ISBN: 978-1610390934.

¹² Nielsen, M. (2012). Reinventing discovery: The new era of networked science, Princeton University Press, ISBN: 978-0-691-14890-8.

their habits. Likewise project Galaxy Zoo has recruited more than 200,000 volunteers to view and classify galaxy images. This process has led to entirely new discoveries about our universe. William T.G.Morton's discovery of surgical anesthetic ether (1846) is the last *individual* medical discovery. Since then all drugs have been discovered by collaborating teams (Kirsch and Ogas (2017)).¹³ Even when we consider early days when lone wolves were more common, there was some implicit collaboration. When they had to add an engine to the Wright Flyer, the brothers turned to Charlie Taylor, a trusted mechanic who also managed their Dayton bicycle shop. For the most part, the brothers had the necessary engineering skills, but they did not hesitate to ask for help when needed. Fast forward to May 1927, when another American, Charles Lindbergh created history again with the first (solo) trans-Atlantic flight from New York to Paris. Lindbergh needed to customize the design of his plane to suit the long-haul flight – basically this meant that he wanted to maximize the amount of fuel he carried while minimizing all unnecessary frills. To build his customized plane, Lindbergh turned to a company called Ryan Airlines in San Diego and collaborated with their chief engineer Donald Hall and Benjamin Franklin (sic) Mahoney, the owner (Berg (1998)).(see 5) Good product managers have a sense for the specific pieces that need outsourcing and they have a talent for assembling the right partners for fostering innovation. If any third-party collaboration is absolutely needed for product success, it must be done mindfully, with costs or profit-sharing negotiated in advance. For instance, many biotech companies such as 23andMe lack the capital needed to fully develop

¹³ Kirsch, D.R. and Ogas, O. (2017). The Drug Hunters, Arcade books. ISBN: 9781628729863 (paperback).

new products on their own – it is common practice for small bio-tech firms to enter into agreements with a larger pharmaceutical company to develop their drug candidates in exchange for profit sharing down the road (e.g., 23andMe had such a partnership with GSK).

Have a plan to recruit the first set of customers

The first customer who expressed any interest at all in the Wright Flyer was the British government, which immediately anticipated the military potential of a flying machine. The brothers, with a true sense of patriotism, went back to the US government and offered to work for them instead. Unfortunately, flying was such a novel concept that bureaucracy took its toll – the benefits were still unclear and it took a lot of negotiation to reach any kind of agreement. Eventually, the first customers that the brothers found were actually in France. Moreover, they could not just provide the machines and walk away, they also had to personally train a few French pilots on their planes. (see 3) Product managers will do well to envision a complete services agreement to support the sale of a physical product. Managers can also build credibility during product launch with a live, large-scale demonstration. On the day of the first flight in 1903, only a handful of people (five to be precise) were at hand to witness history in the making. The brothers then went back to their home in Ohio and continued their experiments, but even after they had proved that powered flight was no longer a dream, very few people took them seriously. Flying then, like colonizing Mars today, seemed so bizarre that no one thought it could actually happen. It was only after Wilbur went to France and conducted large-scale demonstrations in front of huge crowds at Le Mans that they achieved credibility. It was a more innocent time and intellectual property might have been shared more openly than today. After the patenting of the Wright Flyer, even the brothers had to defend their turf and wade through a number of lawsuits to fight the competition off – so the demonstration at Le Mans might have

given some aspects of the game away, but it helped to build a cult following and gave them a large share of the market. Product espionage, to be discussed later, was a concern even in those times. Wilbur Wright closely guarded the Wright Flyer and slept with it in a garage during his time at Le Mans (McCullough (2016)).(see 3).

Making flexible products that can adapt to seasonal demand

Airlines usually experience their peak demand during summers – even the end-of-year holiday travel is usually insufficient to tilt the scales (Bouwer et al. (2024)).¹⁴ This seasonality is particularly prevalent in Europe. Given the capital-intensive nature of the airline industry, reducing the flight schedule during the winter months may lead to idling expensive equipment. To adapt, airlines may choose to operate with a lower base level of capacity and ramp up during the summer months by adding capacity via leasing. Aircraft leasing is also studied in Guzhva et al. (2019) who propose this as an excellent tool to manage the balance sheet and financial structure of airlines.¹⁵ Other seasonal strategies identified in Bouwer et al. (2024) include

¹⁴ Bouwer, J, Hausmann, L, Lind, N, Verstreken, C and Xanthopoulos, S. (2024). How airlines can handle busier summers—and comparatively quiet winters, McKinsey Quarterly, January 2024.

¹⁵ Guzhva, V.S., Raghavan, Sunder and D’Agostino, D.J. (2019). Aircraft leasing and financing: Tools for success in international aircraft acquisition and management, Elsevier, ISBN: 978-0-12-815285-0.

identifying countercyclical destinations to help shore up off-season weakness and seeking out alternative demand sources such as commercial and MICE (meetings, incentives, conferences, and exhibitions) contracts.(see 14)

In another interesting example of the impact of seasonality from the world of sports, Muller (2015) rues the fact that international cricket was scheduled in Bangladesh during the monsoons that year, when rain could erase an entire match (and thereby the product offering).¹⁶ Solutions debated to save the product range from roofed stadiums (destroys the cricketing experience a bit and also very expensive) to scheduling matches at a “neutral” country such as the United Arab Emirates during the monsoon season. With the latter approach, the home court advantage will be lost, but the product can still be monetized and matches completed.

One of the cleverest technological innovations to emerge from the challenge of tackling time varying demand are pricing and revenue management systems. Koushik et al. (2012) describe an application of revenue management in the hotel industry.¹⁷ Dynamic pricing enables movie

¹⁶ Muller, A (2015). Cricket in Monsoon Season: World Cricket's Clown of the Month, July 2015, downloaded from <https://bleacherreport.com/articles/2537439-cricket-in-monsoon-season-world-cricket-clown-of-the-month-july-2015> on February 21, 2024.

¹⁷ Koushik, D, Higbie, J.A., Eister, C. (2012). Retail price optimization at InterContinental Hotels group, *Interfaces*, Vol. 42, No. 1, January–February 2012, pp. 45–57.

theaters to lower admission rates on Tuesdays when demand typically might be low and jack up prices during Saturday evening slots. Product managers must decide whether some kind of revenue management system can increase the ROI from their products.

Taking advantage of fractional ownership for divisible products

The average per capita income of customers in rural Bangladesh was at one time less than \$200/year (Prahalad and Hammond (2002)).¹⁸ Despite this low number, these customers are avid consumers of telecommunication services marketed by Grameen Telecom and a village may contribute an average revenue of \$90/month. Grameen Telecom's phones may be owned by a single entrepreneur with access to capital, but the phones are shared by the community, i.e., the product is smoothly divisible, enabling fractional ownership.

Not all product categories are amenable to such fractional ownership. In aviation, we have examples such as NetJets – at one time, the legendary investor Warren Buffet has been an owner of this company and it is clear from his views expressed in his annual letters to shareholders of Berkshire that this was a bumpy ride as an investment. Whyte (2018) describes how Mr. Buffet described the NetJets investment variously as: “I failed you in letting NetJets descend into this condition” in the 2009 newsletter and in 2011 “Without Berkshire's support, NetJets would have gone broke.”¹⁹ There were good years as well: e.g., in 2006: “NetJets is now operating

¹⁸ Prahalad, C.K. and Hammond, A. (2002). Serving the world's poor, profitably. Harvard Business Review, 2002.

¹⁹ Whyte, A. (2018). Buffet on 20 years of owning NetJets: from “huge potential” to “I failed

profitably in Europe.” Overall, in his stellar career, Mr. Buffet has not been a strong endorser of investing in the aviation industry. In a now famous speech made in Sun Valley, Idaho in 1999 (Schroeder (2008)) Mr. Buffet claimed that an investment in the airline industry would have made almost no money in its entire history.²⁰ He points out that in the early years of aviation between 1919 and 1939 there were some 200 companies fighting for market share. It might be the case that the barriers to entry and the precise safeguarding of intellectual property were somewhat different in that era, but these factors are worth a careful evaluation by all product managers considering an entry into a market.

Managing customer service disasters and the impact of social media on products

In 2008, Dave Carroll, a Canadian musician was traveling on United Airlines to give concerts in Nebraska (see Bernoff and Schadler (2010) for the full story).²¹ He made the mistake of

you” and “These problems are behind us.”, downloaded from

<https://www.corporatejetinvestor.com/news/netjets-in-berkshire-hathaway-reports-587/>,

February 29, 2024.

²⁰ Schroeder, A. (2008). The snowball: Warren Buffet and the business of life, Bantam books, ISBN: 978-0-553-80509-3.

²¹ Bernoff, J and Schadler, T. (2010). Peer influence analysis chapter 3 in Empowered: Unleash your employees, energize your customers and transform your business, Harvard Business School press, ISBN-13-978-1-4221-7030-4.

checking in his expensive Taylor guitar – the baggage handlers, likely unaware of the \$3500 price tag, treated his equipment like any other suitcase and tossed it around. The guitar was damaged and to add to Dave’s chagrin, United refused to reimburse him. Dave was miffed and went on to make and post a video titled “United breaks guitars” (available on YouTube at the time of this writing). This open complaint went viral and proved very embarrassing for United. Learning from this mistake, United used the experience to train their customer service representatives.

The notion of tapping customer opinions and expertise to improve products was first proposed by Prahalad and Ramaswamy (2000).²² In recent years, online customer forums have mushroomed, but social media can be a double-edged sword, either providing positive or negative word of mouth. Herhausen et al. (2019) have studied a dataset of nearly half a million negative posts from online forums of 89 companies in the S&P 500.²³ They categorize companies’ responses into five buckets: offering to move the conversation to a private channel, apologizing, expressing empathy, offering compensation or providing an explanation. Of the half a million

²² Prahalad, C.K. and Ramaswamy, V. (2000). Co-opting customer competence, Harvard Business Review, January-February 2000.

²³ Herhausen, D, Ludwig, S, Grewal, D. Wulf, J and Schoegel, M. (2019). Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities, Journal of Marketing, 1-21.

posts, about 3% went viral and many of the viral posts contained intense emotions. One of the important remedies identified is responding *immediately* to the angry comment – ignoring the complaint like United initially did could have dire consequences.

Partnering with labor unions for service-intensive products

Airlines are in the business of transporting passengers whose customer experience relies heavily on the performance of frontline employees such as pilots, crew members and baggage handlers. Many of these employees belong to a collective bargaining unit and the success of the airline and the service product depends upon the quality of partnering with labor unions. Aviation history is replete with stories of nasty confrontations between product managers and labor unions. The most infamous of these is the August, 1981 strike when members of the Professional Air Traffic Controllers Organization (PATCO) walked off the job. President Reagan invoked a 1955 law and declared the strike to be illegal – he gave air traffic controllers 48 hours to think it over and well, return to duty. At the end of the grace period, he fired 11,000 striking air traffic controllers and told them that they could never return – a major blow to the labor union movement in the United States (McCartin (2011)).²⁴ We would imagine that an air-traffic controller strike would bring the nation's air transport system to a grinding halt. But that did not actually happen – Reagan brought in controllers from the military and coaxed some others out of retirement. To boot, the timing of the strike was not judged to a nicety. The day after the strike, skies were blue, weather was good and the substitute controllers had an easy time adjusting. In comparison, in

²⁴ McCartin, J. (2011). Collision Course: Ronald Reagan, the Air Traffic Controllers, and the Strike that Changed America, Oxford University Press. ISBN: 978-0199836789.

more recent times, British airways crew members threatened a strike when the busy summer season was starting (Frost (2022)).²⁵ They were able to win an average 13% pay rise which merely returned their pay to pre-pandemic levels. In 2023, the CEO of Starbucks was forced to write an open letter to unions representing employees urging them to work with management to settle differences (Laxman (2023)).²⁶ Even as this paper was being written, the drama unfolded and the New York Times reported that Brian Niccol would replace Laxman Narasimhan as the CEO.²⁷ Clearly, CEOs, labor unions and service product managers can learn valuable lessons from history. CEOs must serve as a bridge between board members and hard working frontline workers who may feel differently about issues such as wages and hours of operation. Unions will be better served by threatening a strike at a time when it will have maximal impact on disrupting service operations; and product managers must anticipate this timing and put in place stop-gap measures that will enable operations to continue even if skeletally. At the same time,

²⁵ Frost, R. (2022). <https://www.euronews.com/travel/2022/08/16/ba-strikes-will-your-flight-be-affected-and-what-should-you-do-if-it-is-cancelled#:~:text=In%20June%2C%20hundreds%20of%20check,made%20for%20a%20pay%20rise>, downloaded on February 16th, 2024.

²⁶ Laxman (2023). December 11, 2023 letter from Starbucks CEO: Affirming what we stand for; downloaded from <https://stories.starbucks.com/press/2023/letter-from-starbucks-ceo-affirming-what-we-stand-for/> on June 9, 2024.

²⁷ <https://www.nytimes.com/2024/08/13/business/starbucks-ceo-brian-niccol-chipotle.html>

proactive and compassionate communication may be required to convey mutual respect and mutual gain during negotiations. All parties must gauge the environmental mood: strikes were at an all time low in 1982, a year after President Reagan fired the PATCO workers. At the time of this writing, Boeing workers have rejected a management offer and gone on strike and this again reinforces the need to manage workers' needs proactively.

Managing product safety and product recalls

On September 17, 1908, Orville, the younger of the Wright brothers and the first man to conquer powered flight was giving a demonstration at Fort Myer Virginia for the US government.

Although he was very comfortable manning his plane, it crashed and he was hurt badly.

Unfortunately he had another passenger on board: Thomas Selfridge became the first man to be killed in an air crash, creating history.(see 3) Luckily, Orville though badly hurt, survived and lived long enough to see commercial aviation take root.

At the time of this writing, Boeing is being ravaged by various safety concerns with its 737-MAX. On October 29, 2018, a Lion Air plane crashed in Indonesia, killing 189 passengers.

Boeing started a careful investigation, but the FAA did not ground the 737-MAX since it was not clear what had caused the accident – some thought it may have been a pilot error. Tragically, the 737-MAX crashed one more time in short order, on March 10, 2019, near Addis Ababa (Kennedy (2022)).²⁸ This time the accident claimed another 157 lives. After this second

²⁸ Kennedy, R. (2022). Downfall: The case against Boeing, available on Netflix on February 24, 2024.

incident, the 737-MAX was grounded by President Trump and a thorough investigation revealed that the problem may have been caused by an external sensor linking to something called the Maneuver Characteristics Augmentation System (MCAS activation). A congressional hearing was convened under Representative Peter DeFazio and the findings of the commission seem to indicate that inadequate product training, particularly with the MCAS activation system might have been a root cause of the accidents. Boeing was concerned about competition from Airbus, particularly the A320Neo model that was very fuel efficient. In order to accelerate the acceptance of the 737-MAX, the product was marketed without flight simulator training (Kennedy (2022)) and this tactic might have been a contributing factor. (see 28) Another theory revolves around organizational culture: when Boeing merged with McDonnell Douglas in 1997, it purportedly moved from an engineering-focused company to a company focused on financial returns and shareholder value. Indeed, the headquarters was moved from Seattle to Chicago and cost-cutting was prioritized. The lesson for product managers is that engineering personnel must always be given free rein to voice any safety concerns openly, without any fear of retribution. Shareholder return is very important, but must be counterbalanced with safety. Boeing has taken some personnel actions including removing CEO Dennis Muilenberg and the head of the 737-MAX program Ed Clark (Isidore and Muntean (2024)) but recouping the confidence of the traveling public might take a while.²⁹ Airline passengers must believe the old

²⁹ Isidore, C and Muntean, P. (2024). Boeing removes head of 737 Max program in wake of safety incidents, downloaded from <https://www.cnn.com/2024/02/21/business/boeing-removes-head-of-737-max-program-in-wake-of-safety-incidents/index.html> .

saw “*If it ain’t Boeing, I ain’t going*” and to achieve this level of trust, Boeing can showcase its engineering initiatives and quality management programs in the future.

Learning across product generations: The VS300 Helicopter and the V22 Osprey

Consider a contemporary military cargo helicopter, the V22 Osprey. The main selling point of this craft is a tilt-rotor feature – this enables the craft to take off vertically like a helicopter; but after take-off, it can *tilt* its rotors forward and act like a normal airplane, flying at normal aircraft speeds. This provides a great advantage in military situations where an aircraft may have to land in tight spaces, but also have to be able to carry heavy cargo quickly over long distances. The innovations that led to the formulation of this craft took place over several *decades*. The tilt rotor concept was first invented at Bell helicopter by Kenneth Wernicke and Robert Lichten and the first such flight of a tilt rotor craft took place on August 11, 1955, on the Bell XV3-Ship 1. On December 18, 1958, the first 100% tilt rotor operation took place on the Bell XV3-Ship 2. (see <https://vertipedia.vtol.org/biographies/getBiography/biographyID/36>). After several early accidents, the wings that carried the tilt rotors were stabilized for vibration. The design of the Osprey also adopted a twin-tail design for improving safety, but this idea was first invented by engineer Clarence “Kelly” Johnson at Lockheed-Martin. Mr. Johnson designed the Electra 10, an aircraft flown by Amelia Earhart in 1937.

In 1943, engineer Michael Buivid applied for a patent for a feature to fold helicopter blades, making them compact for storage, particularly when transported on big naval vessels. From the Sikorsky “Dragonfly” H5 helicopter, the Osprey V22 adopted the rotor folding mechanism to enable it to be stored in compact spaces. The first lesson is that solutions to a problem with a current aircraft can come from another product designed generations ago, so we should broaden

the scope of our search for a solution to a design problem. Another valuable lesson from this story is that innovators can come from anywhere and smart product managers must grab up human capital as and when they see them. Mr. Buivid was born in Poland and moved to Kiev, Russia where he was befriended by Igor Sikorsky. Mr. Buivid actually served in the Russian armed forces in WW I and immigrated to the United States in 1919, along with Mr. Sikorsky (Sikorsky Archives, October 2012 issue).³⁰ Developing a *knowledge management capability* (Liebowitz (1999)) can be helpful in learning across product generations.³¹

Assessing geo-political risks: Airbus vs. Boeing

The *perceived nation of origin* for any product could affect its branding and overall success in the marketplace. Consumers may have a high regard for “German engineering” in cars whether it be justified or not. The aircraft manufacturing industry is not immune to this national-origin effect and companies such as Boeing and Airbus compete on a global theater fiercely for orders. On August 17, 1993, President Bill Clinton put through a call to King Faud of Saudi Arabia (Lynn (1997)).³² He was lobbying for the Saudis to place a big order with the American aircraft manufacturers (Boeing and McDonnell-Douglas at that time, before they merged). The United States had played a significant role in the liberation of Kuwait in the recent war, and President

³⁰ Sikorsky, I. (2012). *Sikorsky Archives News, October 2012*.

³¹ Liebowitz, J (1999). *Knowledge Management Handbook*, CRC Press, ISBN: 0-8493-0238-2.

³² Lynn, M. (1997). *Birds of prey: Boeing vs. Airbus, A battle for the skies*. Four walls eight windows publishing, ISBN: 1-56858-086-X.

Clinton may have subconsciously expected a *quid pro quo* as a result. Airbus retaliated by unleashing the French president Francois Mitterrand on the Saudis and he tried to win the order for the European company. When the American airframe companies finally won the order, President Clinton claimed a victory lap. Product design for aircraft manufacturing is an incredibly tough business because the airline “must guess what size bucket (size of airplane) is needed to carry an unknown quantity of sand (passengers)” (Newhouse (2008)).³³ At the time of this writing, Airbus seems to have an upper hand because Boeing is reeling from the aftermath of the 737-MAX crashes in 2018 and 2019. However a study of history reveals that this was not always the case. On June 14, 2006, Airbus announced a six to seven month delay in the delivery of the A380 aircraft (Newhouse (2008)).(see 33) As a result of this glitch, operating income was forecast to drop by \$625M/year. In the early years of Airbus, in order to appease the company’s French and German shareholders, the company had a dual management structure, with two CEOs and this may have contributed to the poor running of the company. Another important strategic question is which parts of the complex airframe manufacturing process to outsource: Boeing has outsourced airframe wing manufacturing to Japanese companies (Newhouse (2008), chapter 8). Can these companies turn on Boeing at a later point in time and become Boeing’s direct competitors? At the time of this writing, Chinese companies are experimenting with their COMAC C-919 line – no big orders have been landed outside China but perhaps with some pressure from the Chinese government, domestic airlines may choose the COMAC over Boeing

³³ Newhouse, J. (2008). Boeing versus Airbus: The Inside Story of the Greatest International Competition in Business, Vintage books, ISBN: 978-1400078721.

or Airbus planes. Finally for some market segments, *only* domestic producers fit the bill. A good example consists of military planes: the aviation pioneer Howard Hughes specialized in this market segment and built the gigantic Hercules-H-4 plane to carry military cargo. It made imminent sense as ships carrying military cargo were being targeted by German U-boats during World War II and a solution in the sky may have been called for. At the time of building, the H-4 had the longest wingspan of any aircraft in the world (Hughes (1947)) and was nicknamed the “Spruce Goose”.³⁴

Managing product espionage and product sabotage

Product espionage has even made it into popular culture through the movies we watch, e.g., the 2009 film “Duplicity” that focuses exclusively on corporate spying, or in the 2004 film the “Aviator” in which Howard Hughes accuses Juan Trippe, the legendary CEO of Pan Am of stealing ideas regarding the interior design of his aircraft. One of the historically important figures in aviation is Octave Chanute, a French-born American citizen – the Wright brothers had a Protean relationship with him. On some occasions he acted as mentor and claimed that the brothers were his “pupils.” When the brothers were working at Kitty Hawk, Mr. Chanute sent a couple of observers ostensibly to help them, but at least one of them played interference (McCullough (2016)).(see 3) When the brothers filed suit against a gentleman named Glen Curtiss for patent violation, Octave sided with Mr. Curtiss, infuriating them. Another one of the reasons the brothers encountered initial difficulties was that they were using data published by

³⁴ Hughes, H. (1947). https://en.wikipedia.org/wiki/Hughes_H-4_Hercules, downloaded on March 2, 2024.

another aviation pioneer Otto Lilienthal and they were using this information to design their flight experiments, assuming it was gospel. But for whatever reason, this data was a little off. Finally, the Wrights designed their own wind tunnel experiments in Dayton to recalculate something called a Smeaton coefficient and this adjustment to their data eventually led to success. So if at first your data doesn't work as expected, product managers must create and validate their own data that they trust implicitly. Wang et al. (2023) have recently surveyed the enormous benefits that AI can bring to scientific innovation.(see 10) However, even today, they point out that poor data quality remains an important barrier to innovation. In more recent times, an aviation engineer working for General Electric received what seemed an innocuous invitation from a Chinese university on LinkedIn, in 2017. The university generously offered to pay for his travel and stay in China and more or less flattered him into accepting the invitation (CNBC (2023)).³⁵ While in China, the engineer was allegedly cultivated by Chinese intelligence and recruited to reveal engineering secrets. Stories or incidents such as these may just be fueled by geo-political tensions or political viewpoints and it is quite hard to gauge the veracity of a specific incident or to pass judgments on individuals in a situation such as this. However, it is worth noting that the US-China axis has not always been adversarial. For instance, when Charles Lindbergh was traveling in China, the Yangtze river flooded and Mr. Lindbergh provided invaluable reconnaissance, and this was recognized by General Chiang Kai-Shek who awarded him with the National medal for this assistance. (see 5)

³⁵ CNBC (2023). China's corporate spy war, narrated by Eamon Jabers,
<https://www.cnbc.com/documentaries/spy-war/>

Ensuring product accessibility

Michael Swiatek is the chief strategy officer for the Abra group, a holding company for some Latin American airlines. Since childhood, he has experienced a condition called *retinitis pigmentosa* which leads to night blindness and loss of peripheral vision (Cooper and Fusaro (2024)).³⁶ Mr. Swiatek is a member of the C-suite and has worked for a number of CEOs during his career. He credits his disability with sharpening his listening skills and empathy, much needed qualities for senior managers.

In his interview with McKinsey, Mr. Swiatek outlines how he was able to take a systematic approach to enabling accessibility to airline travel for all. He outlines an audit of the customer journey to develop pain points in four categories: auditory, cognitive, mobility and visual, leading to a catalogue of 90 pain points for passengers. Transporting wheelchairs of passengers is a major hurdle, along with issues like navigating narrow aisles on board the aircraft and an inability to hear airline announcements. Five tools are outlined to ensure product accessibility, including raising general awareness about accessibility, training airline personnel, process reengineering, digital technologies and creating the “hardware” for accessibility. The hardware could include such simple ideas as outlining the flush button in an airplane lavatory with LED lighting to make it easy to spot. Of course all these tools come with a cost, but as Mr. Swiatek states in the interview “transportation is a human right.” He also points out that the IATA

³⁶ Cooper, E, Fusaro, R. (2024). A flight plan for everyone – interview with Michael Swiatek of Abra group, McKinsey Quarterly, January 2024.

conferences now have an accessibility track and managers are beginning to develop KPIs for accessibility, such as the number of complaints registered or even the number of broken wheelchairs.

Keep sprinting because the competition keeps coming relentlessly

The legendary investor Warren Buffet is purported to have said in a speech that he avoids technological innovations such as airlines in his investments. Examining the early history of aviation would seem to vindicate his claim. Orville flew for 12 seconds in 1903. The very same day, Wilbur erased his record, flying for 59 seconds and 852 feet. On July 25, 1909, a Frenchman named Louis Bleriot flew from Calais to Dover across the English channel, claiming a 1000 pound prize. His flight lasted 37 minutes, a dramatic improvement over the 12 second victory six years earlier. In 1911, C.P. Rodgers flew from Long Island to Long beach, California, albeit with eighty stops. In 1913, Roland Garros flew non-stop across the Mediterranean. In 1927, Lindbergh flew solo across the Atlantic. Almost as soon as Lindbergh finished, Clarence Chamberlin and Charles Levine extended his record, flying non-stop to Eislebin, 110 miles short of Berlin, their intended destination. At one point in time Boeing was the major aircraft manufacturer with no competition in sight (alas we must also recall Lotus-123 the spreadsheet software that had 100% market share at one point in time). Airbus has gradually eroded Boeing's market share and even overtaken it in some years.

Creating a fun product innovation culture

Benjamin Franklin's experiments with electricity were imbued with a sense of mischief and fun. He created a charged metal spider that leapt around like a real one and also electrified the fence around his house to amuse visitors (Isaacson (2004)).(see 7) He designed his famous kite experiment to draw electricity from thunderstorms so that he could reach higher than the church

steeple of the time. He also used his own product, the lightning rod, in his own home. But when he connected the rod to earth, he left a small gap in the wiring and installed a bell there so that it would ring whenever electricity passed through. While he thought that this was great fun, his wife Deborah was far from amused as it kept her up when Franklin was traveling (which was often). So he advised her to connect the wire in the gap so that it would protect the house silently (Isaacson (2004)). (see 7) Another advantage of making innovation fun is that we may feel more passionate about what we are having fun with and also take the career risks that product management entails. The Wright brothers were known for their enormous work ethic and were at it six days a week without fail. While only five spectators were present on that historic day in 1903, and despite this lack of a following, a friend, John T. Daniels, was fond of calling the Wright brothers “the workingest boys.” (McCullough (2016)). (see 3) To be the “workingest” of product managers, we must first learn to find joy in our endeavors.

Bridging Academic Learning and Workforce Readiness: The Role and Value of Generative AI and Custom GPTs in Marketing Education

Abstract

Integrating Generative Pre-trained Transformers (GPTs) into marketing education offers a novel approach to bridging theoretical knowledge and practical application. This paper examines the implementation and benefits of utilizing GPT-based simulators in digital marketing curricula as real-life scenarios. By simulating dynamic client interactions, these tools provide students with immersive experiences that enhance critical thinking, adaptability, and readiness for industry challenges. The following paper highlights significant gaps in practical skill development and ethical preparation in traditional education models while demonstrating how these innovative tools help bridge these divides.

Introduction

Integrating Generative AI (GenAI) into academic curricula has become a necessary innovation in the rapidly evolving landscape of education and technology. Traditional teaching methodologies often fail to provide students with the practical, hands-on experiences required for workforce readiness, particularly in dynamic fields like marketing. This gap between academic instruction and industry needs has been emphasized by the Digital Marketing Sector Council (2023), which highlights the importance of aligning educational practices with workforce requirements through collaborations between educators, employers, and job seekers. Similarly, the Society for Human Resource Management (SHRM) underscores that traditional education often emphasizes theoretical knowledge over practical application, necessitating innovative teaching methodologies bridging this divide (Popera, 2024). GenAI tools like custom GPTs allow educators to transform learning environments by simulating real-world marketing scenarios. These simulations create a bridge between academic theory and professional practice, aligning with workforce demands for skills like problem-solving, adaptability, and ethical decision-making.

The Gap Addressed

The primary gap addressed lies in the disconnect between academic business learning and workforce demands (Harvard Project on Workforce, 2023; PLOS ONE, 2023). Traditional educational models emphasize theoretical knowledge, often at the expense of practical application. As industries increasingly prioritize data-driven decision-making and AI proficiency, students must be equipped with both technical skills and ethical reasoning (AI-Enabled ICT Workforce Consortium, 2024; Section, 2024).

Brief Review of Literature

Recent research underscores that 92% of technology roles are evolving due to AI, necessitating reskilling and upskilling to prepare the workforce (AI-Enabled ICT Workforce Consortium, 2024). Furthermore, *The AI Proficiency Report* highlights a significant gap in AI readiness, with only 9% of professionals reporting sufficient proficiency to harness AI's full potential in real-time scenarios

(Section, 2024). Additionally, there is limited research and practical frameworks for integrating GenAI into experiential learning environments. Studies have highlighted that higher education often struggles to align curricula with labor market needs. Research from *PLOS ONE* emphasizes that while education is pivotal in workforce development, significant skill gaps persist between academic training and professional requirements (Chau et al., 2023). Harvard's Project on Workforce similarly stresses the importance of bridging postsecondary education with employment outcomes to better prepare graduates for the workforce (Harvard Project on Workforce, 2023). This paper fills this gap by demonstrating the potential and value of GPTs to:

1. Provide dynamic, real-world simulations for skill-building.
2. Foster critical thinking and adaptability through iterative feedback loops.
3. Enhance ethical awareness surrounding AI technologies.

Significance

A pilot study conducted by Mariani and Powell (2023) underscored the transformative impact of GenAI tools in the classroom. Key findings included:

- **Enhanced Skill Development:** 86% of students reported an improved understanding of marketing responsibilities and expectations. This aligns with the findings by Kong et al. (2021), who emphasize the role of AI tools in enhancing skill acquisition through practical applications.
- **Improved Learning Outcomes:** 84% of participants noted enhanced problem-solving capabilities through practical applications. Similarly, Ding et al. (2024) highlight the role of generative AI tools in encouraging critical thinking.
- **Real-World Engagement:** 73% indicated that AI video simulations created a more immersive and realistic learning environment. Giabbanelli (2023) also demonstrates that AI tools provide compelling, real-world-like scenarios that enhance engagement.

These results highlight the transformative potential and value of GenAI in improving technical proficiency and promoting a deeper understanding of workplace expectations among students. However, the advent of custom GPTs offers an even more profound opportunity for educators. By tailoring GPTs with specific datasets derived from educator expertise, textbooks, open educational resources, and certification materials, instructors can create highly personalized learning experiences. These custom models enable a direct connection to real-world applications, creating value by allowing students to engage with scenarios that mirror actual professional challenges while reinforcing the relevance of their academic learning to industry demands.

Conclusion

This paper underscores the critical role and value of GPT-based simulators in addressing gaps in marketing education by immersing students in practical, real-world experiences. Custom GPTs represent the next evolution, equipping students with technical proficiency, ethical reasoning, and industry-relevant expertise. By leveraging tailored datasets, including instructional materials,

certification programs, and open educational resources, custom GPTs allow educators to craft highly specialized simulations that mirror complex professional challenges. The alignment of these findings with broader research affirms the transformative potential of integrating both GenAI tools and custom GPTs into academic curricula. Future discussions and implementations must prioritize scalability and deeper exploration to fully realize the educational possibilities of custom GPTs in higher education.

References

- AI-Enabled ICT Workforce Consortium. (2024). The Transformational Opportunity of AI on ICT Jobs. Cisco. Retrieved from <https://investor.cisco.com/news/news-details/2024/AI-and-the-Workforce-Industry-Report-Calls-for-Reskilling-and-Upskilling-as-92-Percent-of-Technology-Roles-Evolve/default.aspx>
- Chau H, Bana SH, Bouvier B, Frank MR (2023) Connecting higher education to workplace activities and earnings. *PloS ONE* 18(3): e0282323. <https://doi.org/10.1371/journal.pone.0282323>
- Digital Marketing Sector Council. (2023). Bridging the skills gap in digital marketing. *Canadian Business*. Retrieved from <https://canadianbusiness.com/sponsored/digital-marketing-skills-gap>
- Ding, M., Dong, S., & Grewal, R. (2024). Generative AI and usage in marketing classroom. *Customer Needs and Solutions*, 11, Article 5. <https://doi.org/10.1007/s40547-024-00145-2>
- Kong, J. S. M., Teo, B. S., Lee, Y. J., Pabba, A. B., Lee, E. J. D., & Sng, J. C. G. (2021). Virtual integrated patient: An AI supplementary tool for second-year medical students. *Asia Pacific Scholar*, 6(3), 87–90. <https://doi.org/10.29060/TAPS.2021-6-3/SC2394>
- Giabbanelli, P. J. (2023). GPT-based models meet simulation: How to efficiently use large-scale pre-trained language models across simulation tasks. *Proceedings of the 2023 Winter Simulation Conference*. <https://doi.org/10.48550/arXiv.2306.13679>
- Harvard Project on Workforce. (2023). New research sheds light on the gap between education and employment. Harvard Kennedy School. Retrieved from <https://www.hks.harvard.edu/centers/wiener/news-events/new-research-sheds-light-gap>
- Mariani, R., & Powell, L. (2023). Utilizing Generative AI technology in marketing classes to promote workforce-needed skills. *Presentation of the 2023 Purdue Global Village Conference*. <https://hdl.handle.net/20.500.12264/536>
- Popera, A. (2024). Bridging the skills gap: Preparing future talent for the workforce. *Society for Human Resource Management*. Retrieved from <https://www.shrm.org/enterprise-solutions/insights/bridging-the-skills-gap--preparing-future-talent-for-the-workforce>

Section. (2024). The AI Proficiency Report. Retrieved from https://cdn.prod.website-files.com/635ffd046dcb4346779f7a91/675c4de2b90aa2dcfdca7e71_TheAIProficiencyReport_v02_eBook_v3.pdf

Smart Training for Smart Plants: VR and AI in Wastewater Treatment

By

Olanrewaju I. Ayodele, Graduate Student, Pennsylvania State Harrisburg and

Sofia M. Vidalis, Associate Professor, Pennsylvania State Harrisburg

Abstract

The quality of life in urban settings is closely linked to effective water management, a critical component of smart city initiatives defined by the United Nations as “inclusive, safe, resilient, and sustainable cities.” Wastewater treatment plants (WWTPs) play a vital role in safeguarding public health and the environment by removing pollutants, preventing waterborne diseases, and mitigating nutrient pollution. The adoption of Virtual reality (VR) and Artificial intelligence (AI) technologies in WWTPs has revolutionized operator training and facility management. VR provides immersive simulations for hands-on training, enhancing preparedness and operational efficiency by allowing operators to practice complex procedures in a risk-free environment. AI-driven predictive analytics optimize processes by offering real-time insights into flow rates, chemical concentrations, and equipment performance, enabling proactive maintenance and reducing downtime. The synergy of VR and AI creates powerful tools for training and operational efficiency, providing real-time feedback and personalized learning experiences. This research evaluates the use of VR and AI in WWTPs, highlighting their benefits in enhancing training efficiency, promoting safe practices, and addressing the complexities of wastewater management. Case studies demonstrate the successful implementation of these technologies, showcasing significant improvements in operational efficiency and cost savings. Despite the advancements, gaps in research remain, including scalability issues, integration challenges, and the need for specialized training. Addressing these gaps will further enhance the effectiveness and adoption of VR and AI in wastewater treatment.

Introduction

The quality of life in urban settings is directly tied to water management, as it is a vital resource shared by all cities. Since water management greatly influences other urban services and how they are managed. Smart water management, as part of smart cities projects, offers communities that face concerns including water scarcity, declining water quality, and safety and security issues, all of which are made worse by aging infrastructure, underinvestment, urbanization, and climate change, several advantages if successfully implemented. (Wang, 2025) .

WWTPs play a crucial role in protecting public health and preserving the environment. By removing harmful pollutants from wastewater, these facilities prevent the spread of waterborne diseases and reduce the risk of contamination in drinking water sources (Avantor, 2025). Effective wastewater treatment also helps to mitigate the impact of nutrient pollution, which can lead to

eutrophication and the creation of “dead zones” in aquatic ecosystems (Water and Wastewater, 2025).

Additionally, WWTPs contribute to sustainable development by conserving water resources and supporting climate resilience through the adoption of green energy solutions and innovative treatment processes (Gregson, 2024). Overall, the importance of WWTPs extends beyond immediate health benefits, encompassing broader environmental and economic advantages that are essential for maintaining the delicate balance of our ecosystems (NuWater, 2023). VR technology is being increasingly adopted in WWTPs for training purposes. By creating immersive simulations, VR allows operators to practice complex procedures in a controlled, risk-free environment. This hands-on experience helps bridge the gap between theoretical knowledge and practical application, leading to better preparedness and operational efficiency. For instance, VR can simulate emergency scenarios, enabling operators to practice and refine their responses without the risk of real-world consequences

AI technologies are being utilized to optimize various processes within WWTPs. AI-driven predictive analytics platforms collect and analyze data in real time, offering insights into flow rates, chemical concentrations, and equipment performance. This real-time monitoring allows for proactive maintenance, reducing downtime and extending the lifespan of equipment. Additionally, AI algorithms can predict system failures, enabling timely interventions that prevent costly disruptions. AI and VR are not only two technologies together; rather, they are a synergy that is opening previously unheard-of opportunities in several industries (Baba, 2024).

The combination of VR and AI provides a powerful tool for enhancing training and operational efficiency in WWTPs. AI-driven analytics can be integrated into VR training programs to provide real-time feedback and performance assessments. This integration allows trainees to receive immediate, personalized feedback, helping them to quickly identify and correct mistakes. The use of AI in VR simulations also enables the creation of more realistic and dynamic training scenarios, further enhancing the learning experience.

A wastewater treatment plant’s operational efficiency, troubleshooting abilities, sustainability practices, clean water discharge, environmental protection, public health safety, cost-effectiveness, and regulatory compliance are all dependent on competent operators who comprehend the intricate processes involved in wastewater treatment operations, which is why effective training is essential. Some effective training methods in wastewater used today are as follows:

- **Classroom Instruction:** Providing theoretical knowledge about wastewater treatment processes (Pa. DEP, 2016).
- **Hands-On Training:** Allowing operators to work directly with equipment and systems (Carlton, 2024).
- **Simulation-Based Training:** Using computer simulations to replicate real-world scenarios (Carlton, 2024).

- **Workshops and Seminars:** Offering opportunities for continuous learning and skill enhancement (Carlton, 2024).
- **Online Courses and Webinars:** Providing flexible learning options for operators to study at their own pace (Carlton, 2024).
- **Mentorship Programs:** Pairing new operators with experienced mentors for guidance and support (Pa. DEP, 2016).

Objective

The objective of this paper is to thoroughly evaluate the implementation of VR and AI in enhancing the training efficiency of wastewater treatment plant operations. As the industry navigates increasing complexities due to regulatory pressures, technological advancements, and environmental concerns, the adoption of effective training methods becomes paramount. This study aims to explore how VR can provide immersive learning experiences that deepen the understanding of Wastewater Treatment Plant (WWTP) operations, promote safe working practices, and streamline training processes. Additionally, it will examine how AI can be integrated into this training program to offer adaptive learning experiences, real-time feedback, and predictive analytics. By analyzing trainees' performance data, AI algorithms can identify areas for improvement and customize training modules accordingly. This personalized approach ensures that each operator receives training tailored to their specific needs and learning pace, ultimately enhancing the overall efficiency and safety of WWTP operations.

Applications of VR and AI in WWTPS

VR and AI technologies in wastewater treatment plants have revolutionized the way operators are trained and how facilities are managed. These advanced technologies offer innovative solutions to enhance operational efficiency, improve safety standards, and ensure regulatory compliance. By leveraging the immersive capabilities of VR and the analytical power of AI, WWTPs can provide more effective and engaging training experiences, optimize processes, and address the growing complexities of wastewater management. A study was conducted to identify various wastewater treatment plants around the world that use VR and/or AI in their training programs and the tools they employ. Table 1 summarizes these tools and their benefits for training.

The tools shown in Table 1 have many benefits for training. Mimic Field 3D uses existing 3D CAD models or laser scans of areas to create an immersive and interactive plant infrastructure. It allows operators to practice and become more familiar with operations and maintenance before the plant is built. It also enables training in a safer environment, which reduces the risk of accidents during actual operations (NECI, 2025). Custom VR Solutions is another type of immersive learning tool that provides hands-on experience by simulating real-world scenarios (Alpha, 2024). This tool reduces the need for physical training setups and downtime on operational equipment (NNTC, 2021).

EON Reality VR simulator is a little more advanced, which allows the simulation of both the process control system and primary components of a wastewater treatment plant. There is a mobile version of this tool that allows the operator to train anywhere (FESTO, 2025). In addition, it combines various disciplines such as mechanics, electronics, chemistry, and biology. Custom VR tours offer an immersive tour of the plant so that trainees understand the layout and operations (VRM, 2025). Some options of videos, quizzes, and more, enhance the retention of trainees through an immersive experience. Another VR program is EON-XR. EON-XR is a platform designed to create, share, and experience immersive AR and VR content. Some key features that this program offers are 3D and 360-degree platform content creation, cross-device compatibility, interactive and customization learning for educational and training purposes, and accessibility for a larger audience. EON-XR is widely used in various fields, including education, training, and industrial applications, to enhance learning and operational efficiency through immersive technology (EON-XR, 2025). Figure 1 shows a VR screenshot from EON-XR on a wastewater treatment plant with educational tools such as quizzes, videos, and a popup screen to ask any questions to the AI assistant.

Table 1: Global Wastewater Treatment Plants Leveraging VR and AI for Training

Wastewater Treatment Plant	Location	Uses VR in Training	VR Training Tools	Uses AI in Training	AI Training Tools
Alpha Wastewater	California	Yes	Mimic Field 3D	No	N/A
Emerson Wastewater	Texas	Yes	Custom VR Solutions	Yes	Predictive Analytics
Veolia, Wellington Moa Point Wastewater	Branches in Europe, America, New Zealand, & Australia	Yes	Custom VR Solutions	Yes	AI-powered Monitoring system
Festo Didactic	New Jersey	Yes	EON Reality VR Simulator	No	N/A
Seafield WWTP	Scotland	Yes	Oculus Rift	No	N/A

Melbourne Water	Australia	Yes	HTC Vive	No	N/A
Atotonilco WWTP	Mexico	Yes	Custom VR Tours	No	N/A
Torre Vieja Desalination Plant	Spain	Yes	Custom VR Tours	No	N/A



Figure 1: EON-XR VR Wastewater Treatment Plant

Virtual Reality headsets, such as the Oculus Rift and HTC Vive, are being utilized alongside various software programs to create a safer and more effective training environment for wastewater treatment plant operators. These advanced tools help reduce the risk of accidents by allowing trainees to practice emergency procedures and routine tasks in a controlled, virtual setting. Trainees receive real-time feedback, which enhances their learning experience and accelerates the training

process. Additionally, VR training scenarios can significantly reduce the costs associated with traditional training methods.

This innovative approach is particularly beneficial for secondary treatment processes in wastewater treatment facilities in Pennsylvania. The state's most polluted water, Acid Mine Drainage (AMD) wastewater, contains high levels of dissolved irons (both ferrous and ferric), sulfate ions, and various heavy metals such as manganese, aluminum, copper, and zinc. These contaminants pose significant challenges to the biological treatment processes, as they hinder the growth of microbes responsible for breaking down organic matter.

By integrating an AI-powered monitoring system, facilities can identify early signs of microbe development hindrance, allowing for timely interventions. This proactive approach ensures the efficiency and effectiveness of the treatment process, ultimately leading to better environmental outcomes and improved water quality.

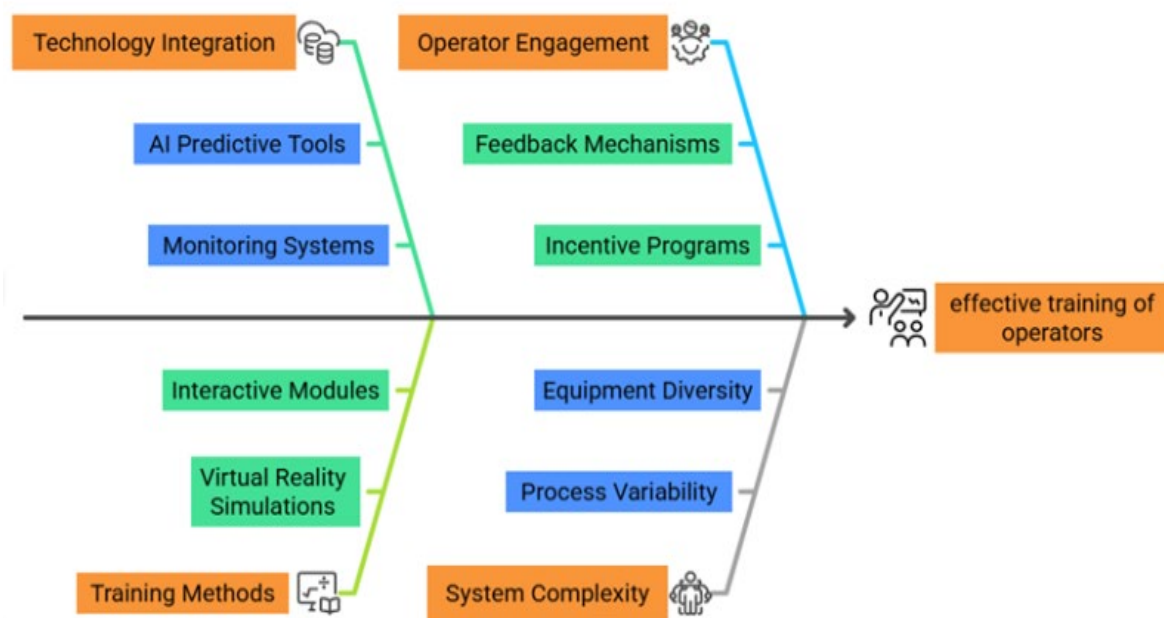


Figure 2: AI Integration in Wastewater Treatment Plants

Case Studies of Current Use

The Emerson Wastewater Treatment Facility in Austin, Texas, unveiled Mimic Field 3D, an immersive training program intended to provide new hires with a thorough grasp of how field changes affect industrial plant operations. With the aid of this VR-based solution, personnel can practice appropriate procedures before visiting potentially dangerous plant locations and gain simulated hands-on experience to better prepare for any plant incident. “The way we train the

workforce of the future is being transformed by Emerson's digital twin portfolio. To close the widening skills gap and boost training efficacy, our industry will depend more and more on VR technologies like these", according to Jim Nyquist, group president of systems and solutions at Emerson (Emerson, 2020).

In a 2018 interview, Scott McMillan, innovation and safety manager at Melbourne Water, stated, "We only identified six design issues using the traditional method, we discovered 20 in VR, which allows viewers to explore the plant", the risks noted could appear minor, such as windows in the control room that overlook tank systems or less-than-ideal valve heights (steel, 2018). However, McMillan explains how minor design flaws that designers fail to notice in the early stages can become serious problems for plant staff members' safety and health at work. "These problems can affect plant operation and maintenance in the long run, in addition to posing risks," he continues. We can see that a VR viewpoint is a huge advancement for design assessment.

A VR simulator for water and wastewater has been developed by Festo Didactic Wastewater Treatment Plant and EON Reality. It includes many scenarios where users may control machines, interact with a virtual water treatment facility, and complete emergency protocols. Users can experience risky and hard-to-replicate events firsthand through VR (Manalac Alvin, 2020).

A significant water utility was dealing with expensive repairs and service interruptions due to frequent pump failures. The utility chose to collaborate with Wetwork Group to deploy a predictive maintenance system powered by AI. The AI systems were able to spot the early warning indications of approaching pump failures by examining sensor data and previous maintenance logs. Equipped with this knowledge, the utility's maintenance teams could plan focused repairs, frequently before the pumps even display any obvious symptoms of malfunction. The effect was immediate; the utility's operating expenses drastically decreased, and the frequency of pump failures fell by more than 70% (Alpha, 2024).

Another pivotal Scenario worked for a sizable municipal water department that was having trouble with recurring algal blooms in its reservoirs, which were lowering the quality of the water and giving its treatment crews a hard time. To identify the early indicators of algae development, there was an implementation of the AI-powered monitoring system that could examine satellite photos, weather patterns, and water quality data. Equipped with this knowledge, the utility's operators might modify flow rates, treatment procedures, and even initiate focused interventions to stop the blooms from becoming out of control. The outcomes were astounding: the utility's operating efficiency increased dramatically, and the frequency and intensity of algal episodes decreased by more than 80% (Alpha, 2024).

Veolia Australia and New Zealand is pioneering the use of VR technology at the Wellington Moa Point Wastewater Treatment Plant, marking a new era in routine asset inspections. This advanced technology allows employees to access real-time operational information, enhancing both efficiency and safety. Serving 188,000 people, the Wellington Moa Wastewater Treatment Plant is the largest in the area and operates under some of the strictest odor consent regulations in the country. With VR technology, Veolia can quickly identify complex issues, reducing wait times and

increasing productivity. Roger Dunn, Veolia's National Operations Manager for New Zealand, explains: "Our team can review and assess conditions from the safety of the office by lowering a VR camera directly into the treatment plant's assets. Engineers can use their smartphones inserted into a VR headset to get a fully immersive experience." This innovative solution sets a new standard for asset grading, particularly for areas that are difficult to access and often require high-risk confined space entry. Roger adds, "VR is giving Veolia new possibilities, allowing us to 'see' into places that we've traditionally not had access to" (Veolia, 2017).

Also, Veolia advances the use of AI to promote sustainable water management. Data analytics, supervision, optimization, and predictive systems for water, waste, energy, and other environmental resources are offered by Hubgrade, Veolia's distinctive suite of digital services. Provides customized solutions for unique ecological transformation journey, whether it be decarbonization, depollution, or resource saving and regeneration, by using the power of data and the most cutting-edge AI technology in conjunction with our global operational knowledge. Over one-billion cubic meters of water are used annually by industrial processes worldwide, thus businesses are searching for methods to maximize water use and hasten ecological change. By integrating real-time, in-field data, AI, and top-tier subject expertise into a single resource, Hubgrade Water Footprint provides complete insight into water operations, including true water cost, water use ratio, and water balance. According to Bernie Anger, Chief Digital Officer of Veolia Water Technologies & Solutions, "traditional water audits only create a snapshot in time of water consumption at an industrial facility." "Hubgrade Water Footprint goes beyond this steady-state representation of total cost of water and captures real-time emergence of water wastage events at a granular level to systematically identify and eliminate water inefficiency in a way that a traditional water audit cannot." Using AI, the platform generates actionable insights that can be used, for instance, to identify instances of water waste in real time, facilitating efficient water risk management and offering suggestions to maximize the impact of energy consumption related to water circuits, thereby lowering the facility's overall carbon emissions (Veolia, 2024).

Gaps in Current Research

There is currently limited data use in most AI applications in wastewater treatment. They are based on small-scale studies which may not be representative of real-world conditions. The maintenance of sensors and other instrumentation required for AI applications can be labor-intensive and costly. In addition, there has been a shortage of expertise in designing AI systems specifically tailored for wastewater treatment processes. Moreover, AI systems can sometimes cause instability in control loops, leading to suboptimal performance (Capodaglio, 2025).

Current VR applications are often limited to pilot projects and have not been widely scaled up for broader use. There have also been challenges in adopting VR equipment for specialized training because of cost, accessibility issues, and user acceptance of new technologies. Most VR applications are focused on training and hazard identification, with fewer applications in operational optimization and real-time monitoring (Fues, 2023).

Addressing these gaps could significantly enhance the effectiveness and adoption of AI and VR technologies in wastewater treatment plants.

Conclusions and Recommendations

The integration of VR and AI technologies in WWTPs has significantly transformed operator training and facility management. These technologies offer innovative solutions to enhance operational efficiency, improve safety standards, and ensure regulatory compliance. By leveraging the immersive capabilities of VR and the analytical power of AI, WWTPs can provide more effective and engaging training experiences, optimize processes, and address the growing complexities of wastewater management.

VR technology allows operators to practice complex procedures in a controlled, risk-free environment, bridging the gap between theoretical knowledge and practical application. This hands-on experience leads to better preparedness and operational efficiency. AI technologies, on the other hand, optimize various processes within WWTPs through real-time data collection and analysis, enabling proactive maintenance and reducing downtime. The synergy of VR and AI creates powerful tools for training and operational efficiency, providing real-time feedback and personalized learning experiences.

Despite the advancements, there are gaps in current research, including limited data use, labor-intensive maintenance of AI systems, and challenges in scaling up VR applications. Addressing these gaps could significantly enhance the effectiveness and adoption of AI and VR technologies in wastewater treatment plants.

Future research will focus on best practices for implementing VR and AI training programs, strategies for overcoming barriers, and further exploration of these technologies' potential to revolutionize wastewater management. In addition, calculating the average adoption rate, cost savings, user engagement, safety improvements, and training effectiveness across all WWTPs can show the strength and direction of each relationship. By doing so, WWTPs can continue to improve their operational efficiency, sustainability practices, and overall contribution to public health and environmental protection.

References

1. Wang, A. J., Li, H., He, Z., Tao, Y., Wang, H., Yang, M., Savic, D., Daigger, G. T., & Ren, N. (2024). "Digital Twins for Wastewater Treatment: A Technical Review," *Engineering*, 36, 21–35. <https://doi.org/10.1016/J.ENG.2024.04.012>, Accessed January 2025.
2. Baba, yigit. (February 1, 2024). "The Impact of Artificial Intelligence on Virtual Reality-Capsule Sight." <https://capsulesight.com/vrglasses/the-impact-of-artificial-intelligence-ai-on-virtual-reality-vr/>, Accessed January 2025.

3. Capodaglio, Andrea G., and Arianna Callegari (January 10, 2025). "Use. Potential, Needs, and Limits of AI in Wastewater Treatment Applications," *Water*, 17(2), 170, <https://www.mdpi.com/2073-4441/17/2/170>, Accessed January 2025).
4. StartUs Insights (2025). "Tope 10 Wastewater Treatment Technologies (2025 & Beyond)," [Wastewater Treatment Technology in 2025 | StartUs Insights](#), Accessed January 2025.
5. Nagpal, Mudita, Miran Ahmad Siddique, Khushi Sharma, Nidhi Sharma, and Ankit Mittal (July 26, 2024). "Optimizing Wastewater Treatment through Artificial Intelligence: Recent Advances and Future Prospects," *Water Science & Technology*, 90(3): 731-757, [Optimizing wastewater treatment through artificial intelligence: recent advances and future prospects | Water Science & Technology | IWA Publishing](#), Accessed January 2025.
6. Fernando, W. Ashane M., Sabeeha N.B.A., and Phaik Eong Poh (2022). "Artificial Intelligence in Wastewater Treatment Systems in the Era of Industry 4.0: A Holistic Review," *Artificial Intelligence and Environmental Sustainability: Challenges and Solutions in the Era of Industry 4.0*, Springer, Chapter 3, pp. 57-97, <file:///E:/Olan/AI%20in%20Environmental%20Sustainability.pdf>, Accessed on January 2025.
7. Yu, Xiaoman, Shuai Chen, Xiaojiao Zhang, Hongcheng Wu, Yaoguang Guo, and Jie Guan (September 22, 2023). "Research Progress of the Artificial Intelligence Application in Wastewater Treatment During 2012–2022: A Bibliometric Analysis," *Water Science & Technology*, <https://iwaponline.com/wst/article/88/7/1750/97652/Research-progress-of-the-artificial-intelligence>, Accessed on January 2025.
8. Pennsylvania Department of Environmental Protection (Pa. DEP) (2016). "Wastewater Operator Certification Training: Introduction to Wastewater Treatment," https://files.dep.state.pa.us/water/bsdw/operatorcertification/TrainingModules/ww01_intro_to_ww_treatment_wb.pdf, Accessed January 2025.
9. Avantor Science Central (2020). "Why is Wastewater Treatment Important?," <https://www.avantorsciences.com/us/en/support/tech-articles/masterflex/importance-of-wastewater-treatment>, Accessed January 2025.
10. Water and Wastewater (2025). "The Importance of Wastewater Treatment: A Critical Component for Sustainable Development," <https://www.waterandwastewater.com/why-is-wastewater-treatment-important/>, Accessed January 2025.
11. Gregson, Clover (April 2024). "Why Waste Water Treatment Matters: Understanding the Importance of Treatment Plants," *Our Better Class: Tech*, <https://ourbetterclass.com/why-waste-water-treatment-matters-understanding-the-importance-of-treatment-plants/>, Accessed January 2025.
12. NuWater (June 28, 2023). "The Importance of Wastewater Treatment: Protecting Our Environment and Health," *Water Treatment News*, <https://nuwater.com/the-importance-of-wastewater-treatment/>, Accessed on January 2025.
13. Carlton, Bobb (January 15, 2024). "Wastewater Management is Improving Thanks to AI Technology," *FSSTUDIO*, <https://fsstudio.com/wastewater-management-is-improving-thanks-to-ai-technology/>, Accessed January 2025.

14. NECI (2025). “DeltaV Mimic Field 3D,” VR product, <https://www.neci.com/products/operations-and-business-management/digital-twin-software/mimic-field-3d/>, Accessed January 2025.
15. Alpha Waste Water (August 2024). “Wastewater Treatment and Virtual Reality Training: Enhancing Operator Competence,” <https://www.alphawastewater.com/wastewater-treatment-and-virtual-reality-training-enhancing-operator-competence/>, Accessed January 2025.
16. NNTC (December 14, 2021). “A Stitch in Time Saves Nine. VR Training to Raise the Water Treatment Company’s Effectiveness and Safety,” Innovative Technology Company, <https://nntc.digital/blog/vr-training-for-water-treatment/>, Accessed January 2025.
17. FESTO (2025). “Mixed Reality: Digital Learning,” https://www.festo.com/us/en/e/technical-education/digital-learning/mixed-reality-id_31287/, Accessed January 2025.
18. Virtual Reality Marketing (VRM) (2025). “VR Waste Management Facility Training,” EON Reality, <https://www.virtualrealitymarketing.com/case-studies/vr-waste-management-facility-training/>, Accessed January 2025.
19. EON-XR (2025). “VR Software Program,” <https://core.eion-xr.com/>, Accessed January 2025.
20. Emerson (February 6, 2020). “Emerson’s New Virtual Reality Simulation Improves Workforce Safety and Speeds Training,” Emerson US, <https://www.emerson.com/en-us/news/automation/2002-mimic-virtual-reality>, Accessed January 2025.
21. Manalac Alvin (June 11), 2020.” *VR Waste Management Facility Training – Virtual Reality Marketing*,” <https://www.virtualrealitymarketing.com/case-studies/vr-waste-management-facility-training/>, Accessed January 2025.
22. Fues, Jordan (June 8, 2023). “How Augmented and Virtual Reality Enhance the Water and Wastewater Industries,” EnviroSight, <https://blog.envirosight.com/how-augmented-and-virtual-reality-enhance-the-water-and-wastewater-industries>, Accessed January 2025.
23. Veolia (September 28, 2017). “Virtual Reality: Changing the Game of Wastewater Management”, Veolia Australia and New Zealand, <https://www.anz.veolia.com/newsroom/newsroom/virtual-reality-changing-game-wastewater-management>, Accessed January 2025.
24. Veolia. (2024, July 31). *Veolia takes artificial intelligence deployment to the next level to drive sustainable water management | Veolia WTS.* <https://www.watertechnologies.com/about-us/newsroom/veolia-takes-artificial-intelligence-deployment-next-level-drive-sustainable>, Accessed January 2025.

Protecting Genetic Privacy: Lessons from the 23andMe Data Breach

Amelia Olsson, Shaoping Zhao
Stockton University, Galloway, NJ, USA

Abstract:

In 2023, the data breach at 23andMe, a genetic testing and biotechnology company, took the personal and genetic data of 6.9 million of users across the world. While the users DNA records themselves were secure, the hackers gained access to details such as the customer's family trees, birthdays, geographic locations, and passwords. The entire biotechnology industry was shaken because of this. This breach exposed the weak parts of the company's security system and showed the world how difficult it is to protect sensitive personal information in an age and place of lots of cyber security threats. Not only did the incident cause issues within the media, courtrooms, revenue streams, and customer trust, it also created a fallout within the board of directors. In this case study, the severity, cause, effect, and actions to take from this breach are examined. It goes over the need for a well-planned and executed cybersecurity system since the genetic data industry is not protected under HIPAA. By fixing these problems, 23andMe as well as others in the biotech industry can rebuild trust with customers, create more ethical norms and use and share genetic data in a better way.

Introduction: 23andMe was founded in 2006, by Anne Wojcicki, Linda Avery, and Paul Cusenza, and was a very popular biotechnology company. Their headquarters is located in Sunnyvale, California, they have over 800 employees globally. In 2022, before their data breach, they reported an annual revenue of \$299.5 million. As of 2024, they have about 582 employees and reported \$193.16 million in revenue, which is about a 27% decrease from years before. They usually help customers through their online app and website, but the physical locations were where most of the testing and customer service activities were conducted. The company is most well-known for genetic testing and health insights, where they were able to give you an analysis of your health or provide you with a family tree and cultural

background information. For example, based on your saliva, they could give you a list of ethnicities to show that you were 50% Italian, 12% German, 18% Irish, 10% Russian, and 10% Japanese. The data could also show you a detailed family tree and give you connections to people that you are related to, such as a second cousin you may have never spoken to before. Other things they offered were health plans, with suggestions on how to improve your health, including how to eat, exercise, or sleep better.

With over 15 million customers worldwide, they were once valued at \$6 billion, but even before their data breach, they faced issues related to their revenue streams due to competition in the biotech industry. However, 23andMe was successful and popular because it had created technology with direct-to-consumer accessibility, letting users explore their genetics through very simple at-home DNA kits. They would send at-home kits to customers and gather saliva samples that would then be tested at their headquarters to extract all of the genetic and health insights for the customer. “Since its founding in 2006, 23andMe has sold more than 12 million of its DNA kits” (Kindelan, 2024) Also, they partnered with major pharmaceutical companies such as GlaxoSmithKline, and used this data to provide more insight for medical research and drug development within the therapeutic research industry. As 23andMe is in an industry where privacy is very important, their data breach in 2023 came with many financial and customer trust issues.

Current Problems and Challenges: In October 2023, 23andMe reported a large data breach that affected millions of people’s personal information. Hackers were able to get into customers’ accounts through data breaches on other websites using compromised passwords and usernames. Through these accounts, hackers were able to access personal data such as “family trees, birth years, and geographic locations” (Rahman-Jones, 2024). The DNA records themselves were not hacked, which means that hackers were not able to see customers’ sequins of DNA, health risk reports, personalized health reports, raw genetic data, research data, and genetic studies. The data breach happened at a time when 23andMe

was already experiencing some financial challenges related to their business model. After finding their family tree, customers would rarely spend more for the other services they offered.

Customers lost a huge amount of trust in 23andMe's and similar companies' ability to protect their sensitive data, which raised concerns about 23andMe's cybersecurity practices, including weak password security, a lack of two-factor authentication, and more. These issues bring up the need for improvements in the company's information security systems to address current threats and rebuild trust for their customers and partners. In addition to public backlash, the company also experienced internal turbulence when the board of directors resigned after a disagreement with the CEO over the company's future. Also, government agencies in the UK and Canada started a large investigation into whether 23andMe had proper security measures and properly reported the breach. The investigation reinforced the importance of data security within this field because the information can be incorrectly used for surveillance and discrimination if it falls into the wrong hands.

Finally, another large issue presented by the 23andMe data breach was the lack of federal and state laws for genetic data. While sensitive health data is usually protected under the Health Insurance Portability and Accountability Act (HIPAA), 23andMe was exempt from these laws because it technically operates separately from the healthcare field. According to Anya Prince, a law professor who specializes in genetic privacy, "HIPAA does not protect data that's held by direct-to-consumer companies like 23andMe" (Allyn, 2024). Since their information is not protected by HIPAA, it leaves the customers to trust and rely on the policies the company themselves have created, such as internal policies and state-level laws for their privacy protection. Some states like California and Florida, have created laws giving people certain rights over their genetic data allowing them to request to withdraw their DNA samples from 23andMe's databases. However, this process can be difficult and the majority of customers continue to "[opt] into having their genetic data analyzed for medical research." (Allyn, 2024).

Roots of the Problems: The 2023 data breach at 23andMe can be traced back to many issues that may have broken 23andMe even if a security breach had not occurred. These issues are technical vulnerabilities, gaps and loopholes in their privacy rules, and issues on the management level.

One of the largest technical issues was that there was no two-factor authentication for the user accounts. At the time of the breach, the company used only small password protectors, which made most accounts highly susceptible for hackers to hack. This led to credential stuffing attacks, which were the kind of attacks that hackers used to get into 23andMe's customer information. Credential stuffing is "the automatic injection of stolen username and password pairs into website login forms, in order to fraudulently gain access to user accounts." (Mueller, n.d.) Most people use the same or similar usernames and passwords when they log in to websites and create new accounts. "Credential stuffing is one of the most common techniques used to take over user accounts." (Mueller, n.d.) Since credential stuffing is the most common way of hacking, it would have been smart for 23andMe to have better safeguards to protect against this with security solutions such as two-factor authentication.

Another factor that contributed to the breach was that there is no consistent and enforced set of federal laws in the United States that specifically protected the genetic data that companies like 23andMe have. Since most healthcare companies are protected by HIPAA, they are protected by more data security and privacy protections, but because 23andMe is a "direct-to-consumer" company it does not get any protections from HIPAA. This left the company to rely on its own internal policies and varying state laws to create the rules for the way that the company handles people's genetic data.

Another issue was that the company was trying to anonymize data, meaning that they were trying to remove a person's identifying information from their genetic data, for research purposes. While anonymization was created to protect user privacy, it also created situations where customers could not properly remove their data from ongoing studies, because there was no longer a name or identifying information attached to their data. This left the customer feeling powerless, especially after the breach, as they realized their data could still be used and they had no way of stopping this from happening.

23andMe also had so much focus on the growth of the company by bringing in new technology and solving medical mysteries, that they may have taken their attention, time, resources, and funds away from protecting their customers' privacy. Their partnership with GlaxoSmithKline required a lot of time and money, which may have distracted the 23andMe team from their security issues.

Finally, 15 million users having all of their data on one platform created difficulties in managing the scale of the data. When the company started taking more and more customers, they may not have thought about the fact that they would need to strengthen their security to hold information for all of these people and their data. All together, these issues created the perfect storm for 23andMe leaving them easily vulnerable to a data breach.

Revenue Loss: 23andMe is facing significant challenges in trying to stay operational in the genetic testing industry after its data breach. The company recently announced that it had a 40% reduction in its workforce with over 200 jobs being cut. 23andMe's "share price has fallen by more than 70% this year" as the company has also warned investors "of substantial doubt about its ability to continue operating." Their revenue has also been reported to have dropped from \$50 million to \$44 million in the span of 3 months. In a span of a year "losses fell from \$59m to \$75m" (Silva, 2024) Their prices on the stock exchange fell as well. In 2021, on the Nasdaq, its share price was around \$17.65, in November of 2024, it decreased to less than \$5.

Actions Taken by 23andMe: After the 2023 data breach happened, 23andMe created several plans to help address the reasons why data was hacked and to restore customer trust. One of the first steps they took was sending emails to all customers encouraging them to change their passwords. They also realized that they needed a stronger authentication process, so they introduced two-factor authentication. This two-factor authentication added an extra layer of security to make unauthorized access more difficult even if customer credentials were compromised. Since then, 23andMe has been doing regular cybersecurity

audits, upgraded monitoring systems, and began using real-time threat detection tools to spot future data breaches quickly. 23andMe took accountability and quickly informed the people who were affected by the breach. They wrote public statements giving out many details about what happened and why it happened, as well as sharing their plan to try and fix everything. Transparency and accountability helped build customers' trust. 23andMe was also fully cooperative when it came to investigations by the UK's Information Commissioner's Office and Canada's Privacy Commissioner. The company also went over its data-handling practices and updated them to be more similar to privacy laws. As for the class action lawsuits, 23andMe had a \$30 million settlement, which included financial compensation for affected customers and funding for more extensive privacy monitoring services.

New Features to Increase Engagement: In recent months, 23andMe has introduced a new feature called "historical matches." This feature uses DNA technology to identify genetic links to popular historical figures. For example, this feature may be able to show a person if there was a certain group of Vikings in their family, or if they were possibly related to Elvis in some kind of way. Another feature they added in October of 2023 was a Health Action Plan. This feature is available to 23andMe+ members through a subscription. This feature uses health histories and biomarker data to give people personalized health recommendations. These records could be things such as lifestyle changes, adjustments to a person's diet, level of exercise, and sleep. They also offer blood testing and treatment for certain conditions. For example, members can send blood samples for tests related to organ function, calcium levels, medication in the blood and disease detection. By doing this 23andMe has invested in their technology and has added a whole other layer to their business. 23andMe came across some issues when they realized that customers got their DNA tested to see their genetics, and then never spent any more money with the company. These additional features helped bring in more customers, gain back trust, and increase revenue during a hard time.

Additional Measures for Improvement: While 23andMe took a lot of positive steps in addressing the breach, there are additional steps they could have taken. First, the company could look into end-to-end encryption for all genetic and personal data. This makes sure that even if data were to be accessed by hackers, it would stay unreadable to unauthorized people viewing this information. “End-to-end encryption is a method of secure communication that prevents third parties from accessing data while its transferred from one end system or device to another” (Bacon, n.d.) An example of end-to-end encryption would be, Steve writes a message, after he hits send it goes as a encrypted message, he sends this to Alice, once this reaches Alice, she must enter a key in order to decrypt the message and read what this message says. 23andMe could also create a more accessible system when it comes to the withdrawal of data, regardless of state laws. They could let customers withdraw consent and delete their data from the research database whenever they would like, even if they have already anonymized the data. They should also bring in more third-party auditors to conduct frequent detailed security audits. 23andMe could also benefit from creating settings in a user’s profile that lets them manage their privacy settings, including ways they can see who has been accessing their data. Educating customers about cybersecurity, and the best ways to prevent being hacked or to prevent themselves from being part of a data leak, such as not reusing passwords, could also reduce the risk of credential-stuffing attacks in the future. There are also secure websites that let people know when their information has been part of a data leak. Lastly, advocating for strict federal laws on genetic privacy would let 23andMe be a leading example of ethical data management in the biotechnology field. Adding these features allows 23andMe to further protect its user’s information and re-build itself as a trusted company in the biotechnology field.

Conclusion and Recommendations: The data breach at 23andMe is an important reminder on how easy it is to access important information, how important it is for a business to protect its customer’s data, and a reminder on why it is important to change your passwords for different websites and accounts. This breach exposed the personal information of 6.9 million users, which showed that there were many

forgotten aspects, gaps, and issues within the company's cyber security policies and management practices. While 23andMe took immediate action by suggesting password resets, introducing two-factor authentication, being accountable and transparent to their users, and updating their internal monitoring system, these steps alone may not be enough to protect them for long. Although the instability of their board of directors may not be directly related, having a strong board and work team is important for a company if they want to have standards set in place. When all members agree on something, the policies can be a lot stronger, as all of the team is focused on one common goal. The company must keep building on these to ensure long-term data protection and to bring back the trust that customers once had in this company. Moving forward, 23andMe must create a new approach to its data security. First, the company should create stronger measures such as end-to-end encryption and real-time threat detection tools to lower risks and assess them as they occur before they become serious. Continuing to take accountability is important in order to gain back trust. Tools they can use to do this is to allow customers to opt in and out of research that their data would be used in. Third party security audits should also be done frequently. They should also try advocating for more federal laws for genetic data privacy. Beyond physical fixes, to rebuild trust within the company and trust from customers, shareholders, and stakeholders, they need to fix their communication. They also need to educate their customers and workers about the importance of cybersecurity issues, and how to protect against them. Some members of their workforce may be handling and using data in unsafe and wrong ways. Having team meetings and sessions can help workers protect their own as well as customers' information better. Sending out educational information to customers can also help the company. Educating customers on the importance of changing their passwords and opting into the two-factor authorization is important. The company can implement the two-factor authorization, but the customer needs to know how to use it, implement it, and the importance of doing so. The company also needs to publicly show its commitment to ethical practices and customer privacy, possibly through partnerships with more government agencies and watchdogs. This case brings up the importance of organizations properly handling sensitive data in a very technologically

advanced time. For 23andMe and similar companies in similar industries, they need to learn the balance between growing the company and protecting the company and its customers. By learning from this breach and committing to stronger practices, 23andMe has the opportunity to not only recover but to thrive in the biotechnology industry. By making transparency, accountability, and security measures a stronger priority, the company can rebuild trust and create long-term success.

References

23andMe. (n.d.). DNA Genetic Testing & Analysis – 23andMe. Wwww.23andme.com.

<https://www.23andme.com>

23andMe. (2023a). 23andMe Launches Health Action Plan to Provide Personalized Recommendations

Based on Genetic and Other Health Data | 23andMe, Inc. 23andMe,

Inc.<https://investors.23andme.com/news-releases/news-release-details/23andme-launches-health-action-plan-to-provide-personalized>

23andMe. (2023b, October 7). Addressing Data Security Concerns. 23andMe Blog.

<https://blog.23andme.com/articles/addressing-data-security-concerns>

Allyn, B. (2024, October 3). 23andMe is on the brink. What happens to all its DNA data? NPR.

<https://www.npr.org/2024/10/03/g-s1-25795/23andme-data-genetic-dna-privacy>

Imran Rahman-Jones. (2024, June 10). Genetic testing firm 23andMe investigated over hack. Bbc.com;

BBC News. <https://www.bbc.com/news/articles/c3gg17pq5y9o>

Lutkevich, B. (2021, June). What is End-to-End Encryption (E2EE) and How Does it Work? TechTarget.

<https://www.techtarget.com/searchsecurity/definition/end-to-end-encryption-E2EE>

Mueller, N. (2023). Credential stuffing Software Attack | OWASP Foundation. Owasp.org; OWASP.

https://owasp.org/www-community/attacks/Credential_stuffing

Paul, K. (2024, December 2). Why I regret using 23andMe: I gave up my DNA just to find out I'm

British. The Guardian; The Guardian.

<https://www.theguardian.com/technology/2024/nov/30/why-i-regret-using-23andme-i-gave-up-my-dna-just-to-find-out-im-british>

Silva, Joao. (2024, November 12). DNA-testing site 23andme to lay off 40% of its workers. BBC; BBC.

DNA-testing site 23andMe to lay off 40% of its workers

Walrath-Holdridge, M. (2024, September 16). 23andMe agrees to \$30 million settlement over data breach

that affected 6.9 million users. USA TODAY; USA TODAY.

<https://www.usatoday.com/story/money/2024/09/16/23andme-class-action-lawsuit-settlement/75250132007/>

The Mediating Role of Digital Marketing Capabilities on Digital Marketing Performance in SMEs: An RBV perspective

Khiem Dinh Mai

khieemd605@gmail.com

Global Partners Service Center Corp., Taiwan (R.O.C.)

Ja-shen Chen

jchen@saturn.yzu.edu.tw

College of Management, Yuan Ze University, Taiwan (R.O.C.)

Abstract

While there is a lack of consensus regarding the effectiveness of digital marketing provided to customers that companies engage in B2B relationships, the literature appears to give little focus to the influential factors that can improve a firm's digital marketing performance, particularly concerning the role of digital marketing capabilities and the intellectual capital of the firms. Building on a resource-based perspective, this research seeks to propose a research framework and examine how SMEs conduct their marketing campaigns while implementing digital transformation within the organization. The suggested framework can demonstrate how firms execute digital marketing during the digital transformation process by leveraging the benefits of digital technologies. Furthermore, it serves as an investigation to validate essential factors of digital marketing performance, especially intellectual capital in this study, when companies strive for digital transformation. Theoretically, this research enhances the comprehension of the connection between digital transformation and digital marketing capabilities and the effects of intellectual capital on digital transformation by employing digital marketing capabilities to boost digital marketing performance, allowing firms to generate new value and acquire greater competitive advantages in the marketplace. The findings from this research may offer theoretical and practical implications for the utilization and management of digital marketing capabilities in the digital transformation process from a B2B perspective through comprehensive interviews and empirical questionnaire survey.

Keywords: digital transformation, digital marketing, capabilities, intellectual capital, RBV

1. Introduction

Since the late 1980s, strategic issues related to the RBV and enterprise capabilities have attracted increasing attention in the marketing strategy literature. As one of the most influential theories in management, resource-based theory (RBT) was developed to explain the sources of sustainable competitive advantage at the firm level. A fast-changing environment obliges companies to adopt and utilize digital technologies that lead to fundamental and non-reversible change in business processes, affecting the very nature of product-service innovations and competitiveness. In recent years, the Internet has become a key tool enabling new technologies to be applied and adapted to any type of product or service, as well as to communication and

promotion strategies (Leeflang et al., 2014; Mazzucchelli et al., 2021). In this context, business models have been adapted to a new data-era in which digital transformation are configured as a fundamental pillar for the adoption of new technologies that influence key areas of firm behavior. Leveraging digital resources through digital capability to reconfigure business models and value networks is an effective way for companies to deepen open innovation. Advanced digital capability can provide scope for new functionality, higher reliability, greater efficiency, and optimization opportunities that exponentially increase the value which companies deliver to customers. Firms require capabilities to act upon the knowledge resources that they control in order to achieve superior performance in target markets. Those capabilities may be associated with innovation in terms of the successful commercialization of new products are those related to marketing management and the ability of organizations to develop and execute marketing mix decisions to differentiate and facilitate the commercialization of products (Adams et al., 2019). More specifically, this study argues that companies need to develop digital marketing capabilities in the shape of formalized routines that utilize digital resources effectively to ensure competitive advantage. Thus, with the advent of the digital era and the digital transformation of society, businesses are increasingly calling for a greater focus on digitalization in general and digital capability in particular. As such, digital marketing is becoming an increasingly important source of competitive advantage for any organizations (Hagen et al., 2022). Moreover, Erevelles et al. (2016) also emphasized that RBT, which utilized by numerous marketing scholars in recent years, have offered a valuable explanation of digital technology impact on marketing.

With rapid progress of technologies, organizations have become interested in adopting digital marketing practices as this could help to improve the sales efforts of the organizations while digital marketing is concerned with the use of digital channels for promoting brands that are helpful to enhance the sales efforts of the organizations (Rayburn et al., 2021). Pandey et al. (2020) defined digital marketing as the use of internet technologies or activities, which includes internet marketing, digital channels, e-commerce, social media marketing and mobile marketing to achieve the company's objective. As digital marketing activities help to influence the decision-making process of the customers (Degbey & Pelto, 2021), organizations must possess appropriate digital infrastructure so that they can adopt digital marketing practices using various digital means and can promote their brands to the wider audiences. Shams et al. (2024) shows that there are various digital marketing tools and capabilities to improve brand identity, which help in establishing close association with customers to influence their purchasing decisions in this digital era. Digital marketing capabilities (DMC) refers to the skills, tools, strategies, and technologies businesses and individuals use to promote their products or services through digital channels (Mathew & Soliman, 2021). Studies have highlighted that digitalization of business practices can achieve better organizational performance (Chu et al., 2019; Krings et al., 2021). However, our current understanding of the existing literature on what digital marketing capabilities comprise and what the measurements and consequences of digital marketing capabilities constitute is still lacking. In line of this, Hagen et al. (2022) argued that studies investigating the factors that influence the uptake of digital marketing activities are still lacking. In fact, businesses often play the catching-up game with consumers in adopting digital technologies because there is a lack of digital maturity and capabilities, the uncertainty regarding

the performance benefits of digitalization, and the inability to effectively deploy digitalization to support organizational transformation (Wang, 2020; Annarelli et al., 2021; Marsh et al., 2022). While SMEs could improve their performance by adopting digital platforms, Setkute and Dibb (2022) emphasized that they may face barriers to achieving potential benefits. For example, they may be lack of the specific capabilities needed to succeed with digital marketing (Quinton et al., 2018). Furthermore, there is a tendency for these firms to focus on digital marketing as a set of tools, rather than to guide strategy, with the use of social media and digital communication platforms remaining more tactical than strategic. This suggest that digital is being used to support more traditional business practices associated with these firms' emphasis on sales and tactical marketing goals. To address these shortcomings, the present study undertakes an investigation that focuses on digital marketing capabilities and other factors. Both industrial marketing scholars and industrial marketers seek insights to understand how our knowledge and practice of digital marketing has been structured and configured.

Previous research consistently suggests that marketing capability can enhance firms' ability to effectively configure and deploy resources, help build a sustainable competitive advantage, and contribute to firms' revenue and profit growth in the long run. Consequently, determining which marketing capability to develop and how to foster these capabilities has become an increasingly important issue in increasingly competitive marketplaces. Despite its growing popularity, there are existing gaps that this study will attempt to address. First, digital transformation with technological developments are constantly shaping digital marketing in an organization (Melovic et al., 2020). This perspective is further demonstrated by Wu et al. (2024), who examine digital marketing strategy and performance in SMEs. Given this recent development, it is timely to explore how contemporary DMC adapt to these shifts, especially within the context of SMEs. Second, Wu et al. (2022) have figured out that existing studies lack an analysis of the mechanism of digitalization capabilities on open innovation from the perspective of resource integration; and that few studies have considered the impact of internal and external situational factors on the above process, which is the challenge of how to use digital technology to promote open innovation in companies at this stage. Third, in spite of the growing popularity research on digital marketing recently, most of study focus on strategic issues (Olson et al., 2021; Behl et al., 2023; Wu et al., 2024) or in the context of multinational enterprises (Wang, 2020), or illustrating the moderate mediating role of DMC (Hashim et al., 2023). Hagen et al. (2022) indicated that in the broader marketing and management literature, the limited number of studies that focus on the antecedents of the digital marketing activities of individual firms have highlighted the relevance of resources to their digital marketing activities. In line of this argument, Ravat (2024) indicates that in spite of the increasing digitalization of innovation activities is reshaping how marketing organizations practice innovation to date, little research has studied the innovative capability of marketing aspect. While digital marketing research has made strides in bridging the gap between the digitalization and marketing literature, few studies have investigated the competencies and mechanisms required for developing these capabilities, especially within the B2B context and for innovation purposes. Thus, a comprehensive study on the complex relationships between digital transformation, DMC and the digital marketing performance of SMEs in an RBV perspective is still incomplete and needs more investigations.

In a situation that theoretical and empirical evidence on the antecedents of digital marketing capabilities is very scant, previous research and industry reports indicate that firms' resources in the guise of technology, data, organizational capabilities, skills, human resources and other intangible are of paramount importance to guide digital transformation in business (Elia et al., 2021). In light of those knowledge gaps, and building on the RBV theory of the firm, and on the notion of digital marketing capabilities within digital transformation, this study adopts the resource-based perspective as an theoretical foundation and selects SMEs as research subjects for the study in an attempt to answer the research issues: *"How SMEs can achieve advanced development by enabling digital marketing performance through digital marketing capabilities and intellectual capital in the process of digital transformation? And what is the role of external situational factors in this process?"*. By answering this question, our research not only deepens the theoretical understanding of digital marketing but also furnishes marketers with actionable insights for navigating the relationship between digital marketing and performance in a digital era. To do so, this research applies the theory of the resource-based view (RBV), which is suited for analyzing the influence of the heterogeneousness of organizations on their performance as it aims to explain the competitiveness of individual firms by differences in their resources rather than market characteristics. This research hopefully to provide the following contributions. Firstly, we intend to analyze to extract the main dimensional components of digital marketing capabilities (which is the combination of digital capability and marketing capability) and digital marketing performance, which expand the theoretical study of open innovation in the context of digitalization. Secondly, the study deconstructs the mechanisms of digital marketing capabilities affecting firm's digital marketing performance with a resource integration perspective, which fills the gap in studying between digital capability and open innovation in generally in SMEs. Thirdly, this research investigates the role of environmental turbulence, managers' digital awareness and innovation orientation in the process of digital marketing capabilities affecting the process performance. The findings are expected to enrich the study of digital marketing enabled by digital capability, and provide a reference for digital practice and innovation development in SMEs. The remain of this ongoing research consists of section 2 presenting theoretical background, section 3 shows methodology and analysis and discussion and conclusion are placed in the last section.

2. Theoretical Background

Since the late 1980s, strategic issues related to the RBV and enterprise capabilities have attracted increasing attention in the marketing strategy literature. A fast-changing environment obliges companies to adopt and utilize digital technologies that lead to fundamental and non-reversible change in business processes, affecting the very nature of product-service innovations and competitiveness. Leveraging digital resources through digital capability to reconfigure business models and value networks is an effective way for companies to deepen open innovation. Advanced digital capability can provide scope for new functionality, higher reliability, greater efficiency, and optimization opportunities that exponentially increase the value which companies deliver to customers. More specifically, this study argues that companies need to develop digital marketing capabilities in the shape of formalized routines that utilize digital resources effectively to ensure competitive advantage. Thus, with the advent of the digital era and the digital

transformation of society, businesses are increasingly calling for a greater focus on digitalization in general and digital capability in particular.

2.1. Resource-based Theory in Marketing

This research takes a firm-level view, applying resource-based theory (RBT) to investigate the relationship between organizational assets (capabilities and resources) and performance results since this theory has been expanded to address that decision-makers can improve organizational performance. According to Kozlenkova et al. (2014), the resource-based view of the firm (RBV) and the resultant RBT provide an important framework for explaining and predicting the basis of a firm's competitive advantage and performance because firms' intrinsic capabilities and resources that result in competitive advantage are often taken into consideration by RBT. Deng et al. (2024) highlights that the research literature supporting RBT argues that a company's strength stems from its unique resources; while Varadarajan (2024) indicated that firms that have been operating around are the RBT's primary focus as they often have heterogeneous resource positions with respect to different strategic resource inputs based on their evolutionary path, which includes both past actions and the outcomes. Underlining RBT, Shukla et al. (2024) proposed that marketing capabilities play a crucial role as an internal competency. These capabilities enable companies to develop strategies that address the evolving demands of consumers, providing them with a distinctive edge in the market and facilitating swift adjustments to market changes, and thus. Firm performance, efficient use of resources, innovation, and reduction of possible market hazards can be enhanced. Although management relationships provide chances for collaboration, trust, and information sharing, it is marketing skills that guarantee these relationships are translated into strategies that win the markets. Based on RBT, companies are understood by their internal resources and capabilities, such as marketing capability, which fosters cross-functional collaboration and guarantees that innovations are both technically possible and relevant to market demands. These internal resources and capabilities are even more significant than their external industry environment (Collis & Montgomery; 2009). Historically, RBT has been applied most frequently in three domains: marketing strategy, international marketing, and marketing innovation. Haverila and Haverila (2024) stated that resources as such cannot be beneficial unless they enable the firm to create distinctive choices or to implement unique marketing strategies.

2.1. Digital Transformation in Marketing Perspective

Digital transformation is a process of change that leverages the leverage of technology and digital capabilities to create added value through business models, operational processes, and customer experiences. Therefore, digital transformation aims to improve entities by triggering significant changes in their properties through a combination of information technology, computing, communication, and connectivity (Vial, 2019). Westerman et al. (2014) stated that digital transformation is a process to create a company's digital competence, which is determined by seven factors (namely: creating an attractive customer experience, utilizing the power of operational processes, finding new business models, developing a transformation vision, involving employees to make the vision a reality, managing the transformation process, and

building technology leadership capabilities). All of these factors are a process to create a company's digital competence through the development of digital capabilities and leadership capabilities. Businesses' digital transformation is essential for gaining new strategic advantages in the future and advancing high-quality development (Chen & Zhang, 2024)

According to Kannan & Li (2017), digital marketing activities encompass all traditional marketing activities that are enabled by digital technologies. In addition to traditional marketing activities, digital marketing activities are increasingly recognized for their potential to increase the competitiveness. Herhausen et al. (2020) encountered meaningful gaps in many capabilities. Many industrial marketers are being challenged by a deluge of data that is well beyond the capacity of their firms to comprehend and use, as pointed out by the gaps regarding big data infrastructure for machine generated data and unstructured data, big data visualization, artificial intelligence and machine learning, and advanced data analytics. Moreover, the survey revealed multichannel capability gaps, also in terms of providing a mobile app for customers, digital selling tools, and IoT technology. In addition, B2B firms have a capability gap when it comes to taking advantage of positive customer reviews. Setkute and Dibb (2022) found a number of factors were involved in shaping digital marketing practices divided into 2 groups: internal factors (i.e., business culture, capabilities, resources, organizational structure, products) and external (i.e., customers, competition, operating environment and wider trading environment). More specifically, while the internal and external factors underpinning the B2B SME context help explain why digital marketing practices have developed as they have, there will also be variations in the use across B2B SMEs due to their context-laden internal business process, availability of resources and capabilities. Although the industrial marketing literature has often employed the resource-based approach as an important theoretical platform, it is surprisingly scarce in digital marketing research.

2.2. Digital Marketing Capabilities in SMEs

Marketing capabilities (MCs) are defined as the organizational abilities to conduct a set of tasks utilizing the available organizational resources to achieve a desired performance outcome (Herhausen et al., 2020). Generally, for capability theory, a firm can be viewed as a collection of capabilities (Prahalad & Hamel, 1990). The fundamental imperative of capability theory is that, to achieve competitive advantage, and, thereby, superior financial performance, firms should identify, seek, develop, reinforce, maintain, and leverage distinctive capabilities and a capability entails a firm's ability to sustain the coordinated deployment of resources in a way that helps a firm achieve its goals (Madhavaram et al., 2022). This conceptualization suggests a "know-how" that helps the firm achieve its objectives (e.g., developing effective software solutions for client firms).

According to Herhausen et al. (2020), digital marketing capability is the ability of a B2B firm to perform a coordinated set of digital-related tasks (operational or dynamic), utilizing digital resources, for achieving a competitive advantage. Particularly in marketing context, Homburg and Wielgos (2022) defined digital marketing capabilities (DMC) refer to a firm's ability to use digital technology-enabled processes to interact with customers and partners in a targeted,

measurable, and integrated way to create new forms of value without regard for distance or time. Although the industrial marketing literature has often employed the resource-based approach as an important theoretical platform, it is surprisingly scarce in digital marketing research. The author indicated that existing literature reviews focus on published research only, thereby neglecting the accelerating emergent of “new” digital capability in managerial practice. This is troublesome, especially given the speed at which digital technologies evolve and impact business opportunities. The focus of a digital marketing strategy should therefore be on how to use digital resources as inputs which, in turn, create capabilities from which value and competitive advantage can be derived. Consequently, we mainly focus in our review on Digital Marketing Capabilities because these endow firms with the means to adapt to market changes that stem from digitalization, new technologies, and changing consumer behavior. We thus complement existing literature by investigating its effects on the Digital Marketing Performance of SMEs viewing Digital Marketing Capabilities as a combination of two main constructs: Digital capability and Marketing capability.

2.2.1. Digital capability

Firms’ capabilities have been defined as the “*firm’s capacity to deploy resources for a desired end result*” (Helfat & Lieberman, 2002). Extant theorization around firms’ capabilities has defined them as organizational abilities to combine, assemble, integrate, and exploit resources to achieve a competitive advantage (Teece et al., 1997; Eisenhardt & Martin, 2000). Leveraging digital resources through digitalization capabilities to reconfigure business models and value networks is an effective way for companies to deepen open innovation (Rachinger et al., 2018). Nylén and Holmström (2015) defined Digital capability as “*embedded digital capability*” to put in place “*appropriate tools*” for product and service innovation and “*managing new types of digital innovation processes that emerge*”. According to Elia et al. (2021), digital capabilities are defined as the competences by which companies align the features of digital technology with customers’ needs and wants, and which are becoming increasingly relevant for successful digital transformation. Digital capabilities are therefore firm-level capabilities that are inclusive of what has been previously termed as internet capabilities. However, the use of digital technology in today’s companies is still in the exploration stage. Companies are not clear about their digital development paths, and the large investment in digital technology has not yet yielded significant innovation benefits, thus hindering their motivation. Therefore, how to enable open innovation through digitalization capabilities to form a new model of innovation development that adapts to the changing environment has become a strategic issue that many scholars and entrepreneurs focus on in the context of digital transformation (Obradović et al., 2021).

2.2.2. Marketing capability

Jaworski and Lurie (2019) noted that marketing capability can be firm-specific, arguing that there is no single winning marketing planning process. These authors stress that improved marketing capability yield different, higher quality choices than the organization made before, and consequently get better results (p. 373). Therefore, a firm is required to identify its key marketing choices, which are framed by an overarching strategy. Firms should start by building the

capabilities needed to make the three or four most important marketing choices first, then move on to building the capabilities for the five or six next most important marketing choices, and so on. However, in certain settings (e.g., country), such marketing choices can be similar across firms, making some capabilities more valuable in reaching goals. Overall, firms must carefully select which capability is the highest priority based on both internal and external factors. According to Haverila (2024), previous RBT-related research has consistently defined marketing capabilities as *“the integrative process of utilizing firm resources (tangible and intangible) to recognize the specific needs of consumers, achieve competitive product differentiation and realize superior brand equity”* (Day, 1994). More specifically, Wong et al. (2022) defined marketing capabilities at firm-level as a firm’s ability to achieve desired marketing outcomes with its available resources, in which complex bundles of skills and accumulated knowledge are involved that enable firms to cope with market complexity. Research in the literature have identified marketing capabilities as a crucial element that businesses employ to gain a competitive advantage (Najafi-Tavani et al., 2016; more). Since connecting with customers and customer satisfaction are key components of a company’s marketing capabilities, having high-quality, up-to-date market data produced by cutting-edge technology is essential when utilizing current market data to improve customer satisfaction, revenue, and the profitability of the company in the context of digital marketing.

2.3. Intellectual Capital

RBV theory states that identifying and managing intellectual capital is a critical element for better performance. According to Ahmed et al. (2022), intellectual capital is *“the intangible resource of the organization includes skills, knowledge, training or information which helpful for the organization to gain competitive advantage”*. Meijerink and Bondarouk (2018) specifically identified intellectual capital – that is, the knowledge, skills, and abilities a business utilizes in value creation processes – as one of the primary determinants of value creation. Intellectual capital is recognized as an important and vital ingredient for the success of companies in a competitive environment. This construct has been extensively accepted to have a positive influence on firms’ business with evidence from many sectors (Al Queshi et al., 2023). Research suggests that Intellectual Capital can benefit firms through positive influences on its performance in general, that the concept has multiple dimensions while Madhavaram et al. (2022) has shown the constructs in terms of three broad dimensions: human, relational, and structural. In line of this, Al-Omoush et al. (2022) determined that there is a broad consensus in the literature that intellectual capital comprises human capital, structural capital, and social capital. This study adopted the same three dimensions because they are comprehensive and strongly related to the main Intellectual capital components of human, structure and relations. The first dimension, Human Capital (HuC), has been defined by researchers in a similar way, generally containing the same words, reflects employees’ knowledge, capabilities, education level, soft and professional skills, and other personal characteristics (Kucharska, 2022). Structural Capital (StC) refers to processes, systems, procedures, represents the organization’s ability to meet external and internal challenges, reflects an organization’s whole knowledge infrastructure. Structural Capital (or organizational capital) represents the unique knowledge institutionalized and codified by a

firm through policies and procedures, routines, processes, work systems, and management structure (McDowell et al., 2018). The third dimension, Social Capital (SoC), includes the value and knowledge of corporate networks and relationships with customers, suppliers, distributors, competitors and all other related parties (Khattak et al., 2024). More specifically, Agostini et al. (2017) stated that relational capital comprises relationships with external partners, such as customers and suppliers, as well as other relational assets like reputation, brand, and loyalty.

Furthermore, while most typical approach to specify IC is to divide it into three sub-dimensions of human capital, relational (or social) capital and structural (or organizational) capital, recent literature has suggested that renewal capital should be acknowledged as an important facet of IC, as it indicates an organization's capability to survive in changing conditions. According to Kucharska (2022), renewal capital reflects the general organizational ability to learn. So, it seems that renewal capital is the main IC component that influences innovations in an organization, but it is only when supported by other components. Ritala et al. (2023) defined renewal capital as the firm's ability to learn and to renew its knowledge base, addressing the ability of the organization to continuously learn and innovate to sustain, adapt and recreate knowledge. Thus, this study adopted Renewal Capital (RnC) as the fourth dimension in the intellectual capital of a firm aiming for successful digital marketing activities.

The evolution of the field of intellectual capital has coincided with the rise of the knowledge-based economy (Al-Omoush et al., 2022). This intangible capital captures the tacit and explicit knowledge embedded in humans, in organizations, and in relationships with business partners (Mubarik et al., 2021). There is a broad consensus in the literature that intellectual capital comprises human capital, structural capital, and social capital (e.g., Bontis et al., 1999; Han and Li, 2015). Human capital encompasses the knowledge, abilities, educational background, professional skills, and other personal attributes of employees (Kucharska, 2022). Structural capital represents the codified knowledge accumulated in databases, files, manuals, information systems, structures, procedures, routines, trademarks, and organizational culture. It gives a firm's human capital the capacity to learn and innovate (Bontis et al., al.,1999). Social capital is relational: it describes the knowledge rooted in relationships and networks of organizations, including internal and external stakeholders (Han and Li, 2015). It is one of the main sources of information and knowledge gathering for innovation and learning (Hsu and Sabherwal, 2012).

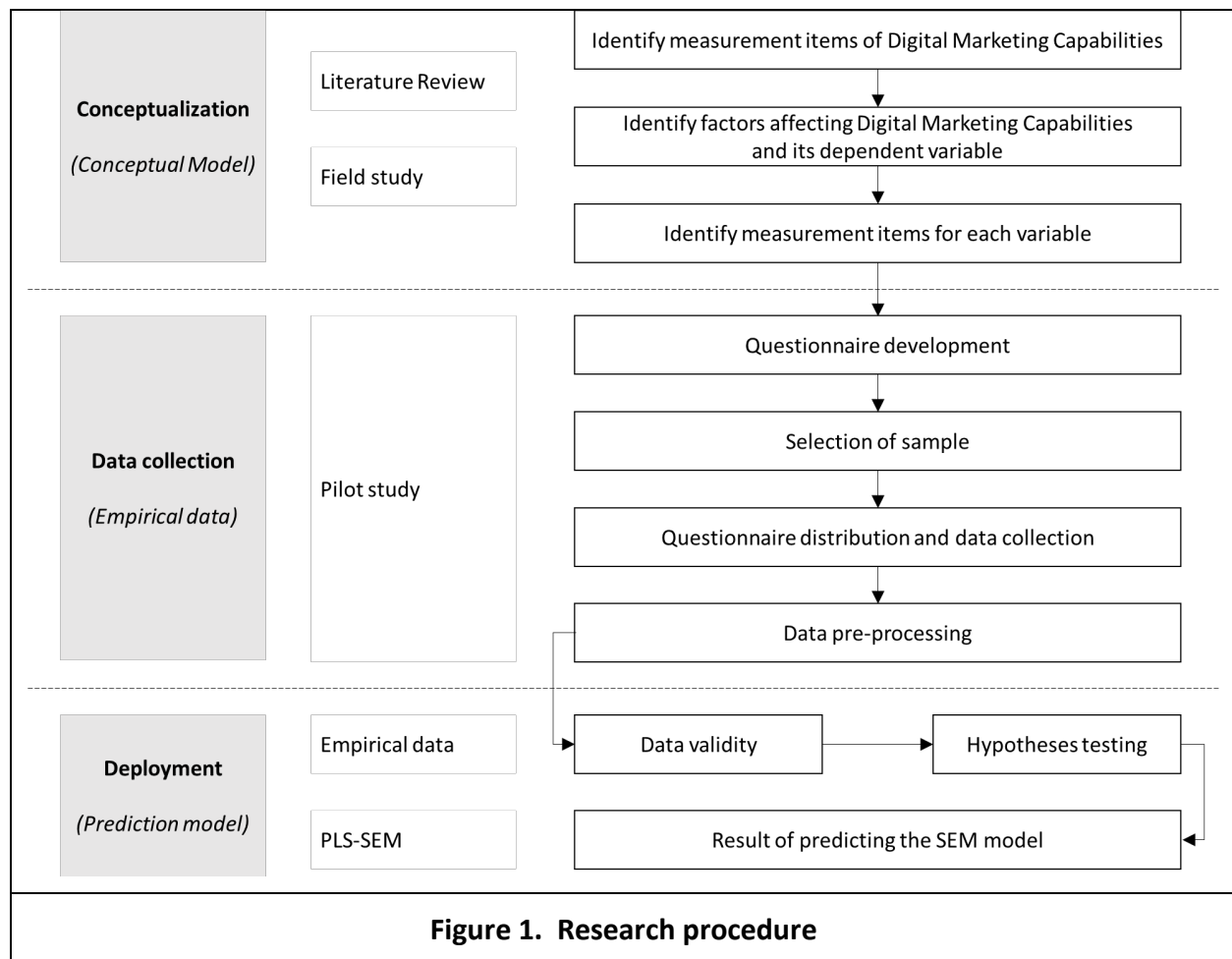
3. Research Method

Given the limited understanding and unclear conclusions regarding how digital transformation impacts on digital marketing capabilities and its performance, and the aim is to broaden the scope of digital marketing research by exploring a new direction, this study delves into the phenomenon through a discovery-driven method, drawing on aspects of grounded theory and a theories-in-use approach (Zeithaml et al., 2020). This study conducts pre-test in-depth interviews with an examination of existing literature review to grasp the intricate effects of digital marketing capabilities from the viewpoints of managers and supervisors, which facilitates the discovery of compelling insights and novel concepts emerging from the experiences of informants, unencumbered by predetermined theories or frameworks. According to Zeithaml et al. (2020)

such an approach involves *“soliciting from study participants ... the ideas they feel are important and how they are linked to one another”* and *“is particularly valuable when scholars want to ... extend extant perspectives and address ambiguities, or guide future empirical efforts”*.

Considering the largely unexplored nature of the impact of digital marketing capabilities on performance, this qualitative approach to investigating digital transformation and digital marketing capabilities also provides some distinct advantages that serves as a valuable way in pinpointing key variables and facilitating comprehension of the relationships between them, while employing a qualitative approach enables us to build upon and contrast with findings from existing research, enrich and broaden our knowledge about the research field (Biemans, 2023).

The study consisted of the following stages: conceptualization, data collection, modeling, and deployment (Fig. 1). The conceptualization phase involved literature review and identifying the measurement items of social coupon redemption decisions. To gain a full understanding of the phenomenon and reflect the qualitative and exploratory nature of the study, semi-structured interviews were undertaken as a field study with a group of 10 practitioners in different SME firms from different industries to elucidate the underlying the issues affecting the adoption and use of digital marketing. Existing networks and connections from the university and the industry were used to approach the participants. After that, an empirical questionnaire survey was conducted to collect data and deploy the prediction model by utilizing PLS-SEM. The data collection phase covered questionnaire development, sample selection, primary survey, and data pre-processing. Finally, in the deployment phase, we evaluated the relationship between the factors and predicted social coupon redemption decisions via the best-fit models, thus addressing research question.



3.1. In-depth interviews

3.1.1. Interview design

To capture a wide range of experiences, a theoretical sampling technique was used that includes informants that are expected to maximize the discovery of variation between concepts. In each interview, there were 2 main aspects were discussed. The first part is about digital transformation in general, while the second one is deeply focused on digital marketing. Before conducting the second part of the interview, we first explained the concept of digital marketing. To avoid the interviewees being limited by the concept of digital marketing, they were encouraged to speak freely about any innovative activity they thought was related to digital marketing.

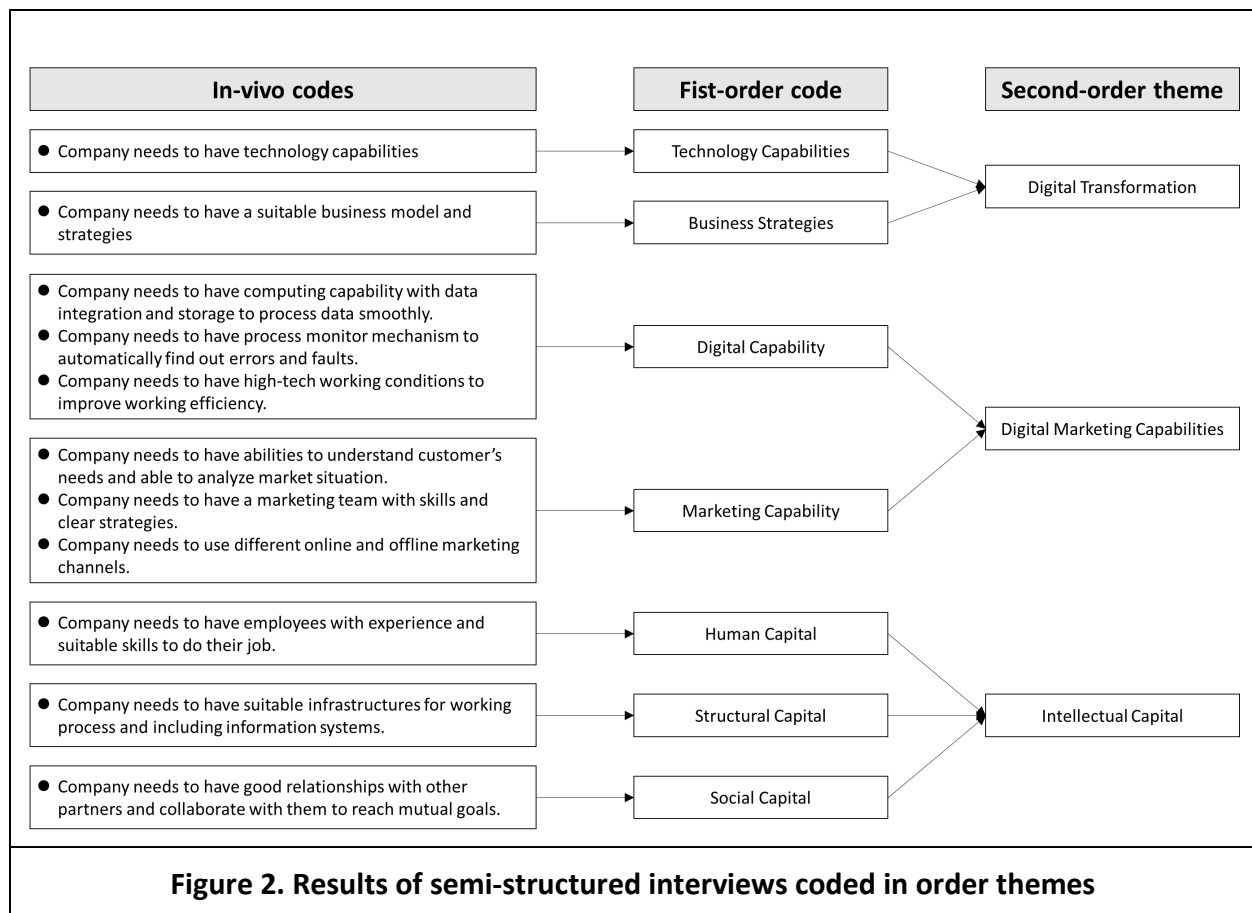
Table 1. In-depth interview question items	
Aspects	Interview question items

Digital Transformation	1. When did the company start its digital transformation? What digitalization attempts have been made in recent years? 2. Which business operations of the enterprise apply digital technology?
Digital Marketing	1. Do your company apply digital marketing into operation? 2. What specific digital marketing activities have companies undertaken in the context of digital transformation? 3. According to you, what are the specific factors affects the integration of digital marketing in your company? 4. What are the main capabilities that the company needs to address in the process of digital marketing? 5. Beside the capabilities, what kind of capitals do the company needs to support for digital marketing? 6. Are there any plans or initiatives of the enterprise in terms of digital marketing in the future?

Table 2. In-depth interview sample details					
Expert ID	Industry	Country	Role	Gender	Interview Length
Expert_01	Education	Vietnam	CEO	Female	15:03
Expert_02	Plastic Production	Vietnam	Vice Head of Manufacturing Department	Male	16:32
Expert_03	Packaging Production	Vietnam	Head of Products Department	Male	14:10
Expert_04	Mechanical Manufacture	Vietnam	Head of Manufacturing Department	Male	20:45
Expert_05	Confectionery (Biscuit and Candies) Manufacture	Vietnam	Director of Marketing Department	Male	15:30
Expert_06	Hand Tools Manufacture	Vietnam	Director Assistant of Manufacturing Department	Male	30:40
Expert_07	Wooden Products Manufacture	Vietnam	Director of Marketing Department	Male	20:10
Expert_08	Construction	Vietnam	Operation Supervisor	Male	15:15
Expert_09	Groceries Retails	Taiwan	Sales Manager	Female	22:30
Expert_10	Groceries Retails	Taiwan	CEO	Female	16:23

3.1.2. Interview data collection and coding

All interviews were recorded and transcribed into text, producing approximately 187 minutes of audio recordings in total. The interview transcripts were initially scrutinized, focusing on remarks highlighted by informants and critical findings about the impact of digital transformation and digital marketing were captured as in-vivo codes illustrating the exact language employed by informants. To be qualified for analysis, the concepts encapsulated in the in-vivo codes had to be explicitly noted by a large number of informants as well as distinctly relevant for many contexts (Ulag & Reinartz, 2011). Following that, axial coding was employed to establish and explore connections between in-vivo codes and first-order categories. The first-order categories were then structured under second-order themes, which form the primary foundations of the findings. The emerging themes were continually refined through ongoing analysis of subsequent interview data. Based on the results of the interview and literature review on academic research, primitive findings were suggested in which eleven first-order categories were identified and five second-order themes (variables) were confirmed as we proposed in the model as shown in Figure 2.



3.2. Conceptual Model Development

3.2.1. Relationship between digital transformation and digital marketing capabilities

A digital marketing capability is the ability of a B2B firm to perform a coordinated set of digital-related tasks (operational or dynamic), utilizing digital resources, for achieving a competitive advantage (Herhausen et al., 2020). As e-commerce started to gain momentum, researchers engaged in revealing and underlying marketing capabilities of successful e-commerce adoption. Furthermore, studies show the importance of the IT infrastructure in determining e-commerce activities. The accompanying Internet technologies have taken fundamental data management software to intelligent, digitized selling, providing the resources needed for customer-centric relationship management so that marketing automation is considered to play an important role in increasing high-quality sales leads and the quality of such resources can substantially impact customer satisfaction and loyalty. The transformation from traditional markets to e-markets demands specialized marketing capabilities housed in a supportive firm that is oriented toward achieving superior marketing efficiency. For industrial sellers there are numerous opportunities to develop specialized capabilities using information technology to enhance communications with buyers and improve distribution, purchasing and supply chain efficiencies. This is especially important for business-to-business (B2B) firms as they often must enter new markets and manage relationships while being separated geographically from their venture markets (Gregory et al., 2019). Furthermore, Ravat et al. (2024) highlighted the necessity for scholars to investigate the key components of digital marketing capabilities as well as its influential factors. The emphasis has been on creative approaches to leverage businesses' use of digital technologies, with a particular focus on improving key performance indicators for their digital marketing initiatives. Through appropriate cooperation between organizations, a more sophisticated and creative approach could be implemented from the necessity to use various forms of digital marketing strategies of other businesses (Sakas et al., 2023). The use of digital technology and supplementary resources to create marketing capabilities that constitute long-lasting relationships with other stakeholders, however, has not received much empirical attention (Cao & Tian, 2020). Therefore, this study mainly focusses in digital marketing capabilities because they can endow firms with the means to adapt to market changes that stem from digitalization, new technologies, and changing consumer behavior.

Hypothesis 1: Digital transformation has positive influence on B2B digital marketing capabilities in SMEs.

3.2.2. Relationship between digital marketing capabilities and digital marketing performance

According to Nordin and Lindbergh (2019), marketing performance refers to outcomes, such as market development, market share, market penetration of new products, business volume, and sales growth; while Paolone et al. (2022) highlighted that marketing performance is the efficiency, effectiveness and adaptiveness of marketing activities with regard to market-related goals, such as revenues, growth and market share. Efficiency, as defined by Morgan et al. (2002), is the ratio of marketing inputs to outputs; while effectiveness, on the other hand, is the degree

to which marketing goals and objectives are achieved; and adaptiveness is the organizational capacity to respond to shifts in the firm's external environment (Järvinen, 2016). Within the context of a digital environment, digital marketing performance (DMP) is defined by EMG (2023) as the degree to which different online marketing strategies, campaigns, and channels are effective in reaching their intended objectives. However, simply investing in resources related to information technologies (IT) doesn't assure success, leading researchers to challenge the direct effect of IT resources on performance. According to Gregory et al. (2019), given the importance to focus on developing resources and capabilities such as e-commerce, one important gap in the literature is the lack of knowledge on how the integration of IT resources and capabilities enhances marketing efficiencies and improves performance. The research in marketing currently integrates both the traditional view with the dynamic capabilities perspective and suggests that marketing capabilities are important predictors of effective marketing strategy, which in turn leads to greater market performance. The competency in evaluating and executing different marketing strategies is needed for continuous adaptation to the market and achieving business success. Implementation of actions when required can increase the operational, financial, and the market performances (Mikalef et al., 2023). With the help of digital technologies, an organization can effectively and quickly make quality decisions than its competitors, which in turn can positively impact organizational performance. Järvinen and Karjaluoto (2015) stated that marketers have realized the value of tracking these interactions and assessing their efficacy as consumers interact with businesses more through digital channels. According to Saura et al. (2021), digital technologies have the power to completely transform the performance of digital marketing by streamlining and automating processes, enhancing targeting and personalization, and enabling greater productivity and cost savings. However, Kozlenkova et al. (2014) suggested that to test the effects of resources (or capability) on performance, research must investigate suitable relevant measures of performance. Notwithstanding that many marketing managers remain dubious about performance measurement and place a greater emphasis on experience and intuition when making decisions, even in spite of numerous marketing studies showing that assessing marketing performance has a positive impact on decision-making. Thus, this study adopts digital marketing performance as dependent variable with following proposition:

Hypothesis 2: Digital marketing capabilities have positive influence on B2B digital marketing performance in SMEs.

3.2.3. Relationship between intellectual capital with digital transformation, digital marketing capabilities and its performance

It is generally acknowledged that intellectual capital is crucial to corporate innovation because it allows businesses to sharpen their competitive advantages and those businesses with greater sums of intellectual capital perform better in terms of innovation. Digitalization has made it possible for organizations to reinvent their processes and embrace more flexible and dynamic methods in an environment of innovative processes. Intellectual capital includes a variety of intangible resources that are vital to the total worth and competitiveness of a company. For the absorption of knowledge and its integration in innovation process the capabilities necessary are required to be acquired. To do so, intellectual capital is a vital competence that equips firms to

deal and manage with the fast-changing innovation landscape. Intellectual capital develops firms' absorptive capacity while supporting the innovation process. Therefore, intellectual capital is a dynamic capability of firms to acquire for the purpose of successful innovation in order to achieve competitive advantage. The human, relational, and structural dimensions of intellectual capital improve tacit and explicit knowledge. This knowledge is related to technology, processes, managerial, and several aspects that lead to develop and enhance the marketing capabilities and digital transformation process, enable firms to utilize the advanced technologies and process of industry 4.0 driven technological transformations in the form of digitalization. It is established that intellectual capital improves the digitalization of firms by managing the capacity that helps in managing and handling such advanced technology and processes (Yang et al., 2024).

Intellectual capital includes three components that complement and support each other. Hence, insufficiency in one of the components leads to an ineffective intellectual capital asset as a whole (Shou et al., 2018). According to Mubarik et al. (2021), organizations with more intellectual capital not only have better human capital and structural capital, but also have excellent social capital, which enables them to collaborate and learn from their business partners. Likewise, Shou et al. (2018) asserted that firms with more intellectual capital are better able to conduct collaborative activities with supply chain partners. Thus, it can be considered as the sum of the entire firm relations that can be used to conduct business and create value. In the context of developing B2B market offerings, Intellectual Capital is expected to make a significant impact and previous research has provided the foundations for the influence of Intellectual Capital on capability in general so the construct is expected to have impact on Digital Marketing Capabilities particularly. Consequently, one particular importance to this research is the notion that Intellectual Capital can benefit SMEs by positively influencing Digital Marketing Capabilities as a basis foundation, as well as Digital Transformation and lead to the enhancement of Digital Marketing Performance. Therefore, it is hypothesized that intellectual capital improves digital transformation, digital marketing capabilities and its performance as propositions are proposed as follows

Hypothesis 3: Intellectual capital has positive influence on digital transformation in SMEs.

Hypothesis 4: Intellectual capital has positive influence on B2B digital marketing capabilities in SMEs.

Hypothesis 5: Intellectual capital has positive influence on B2B digital marketing performance in SMEs.

3.2.4. Environmental Turbulence

Competitive intensity and environmental turbulence are significant contextual factors that impact the performance of firms among various contextual elements. According to Lyu et al. (2020), environmental turbulence is commonly defined as the rate and unpredictable changes in a firm's external environment, in which, technological and market volatility are examples (Hung & Chou, 2013). More specifically, environmental turbulence pertains to the speed and intensity of changes in external factors that are beyond the control of the firm, which might originate from various sources and can take various forms, such as market, technological, or competitive

turbulence, or turbulence resulting from swift modifications in governmental policies, regulations, policies, and natural disasters (Kalubanga & Gudergan, 2022). Environmental turbulences are sporadic disruptions that cause disruptions to the organization and are rarely predictable in their occurrence. The existing studies have emphasized that environmental turbulence compels enterprises to update routinized experiences and methods to improve their ability to respond to environmental changes (Vu & Nguyen, 2022). Since managers frequently perceive turbulence as a threat that could overturn routine actions and past decisions, turbulence has been garnering attention from businesses worldwide.

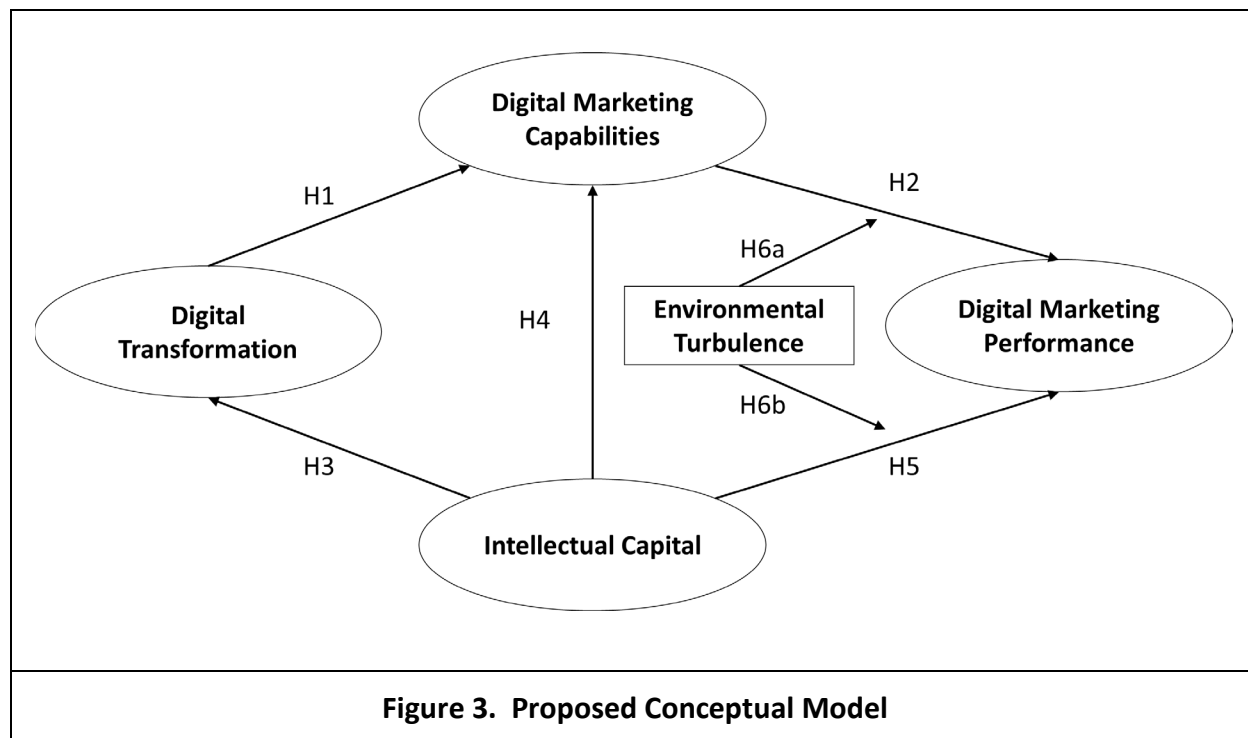
In this study, environmental turbulence is measured using four dimensions that are specifically related to the context of digitalization adopted from the work of Witschel et al. (2022), which are technological turbulence (TT) and market turbulence (MT). Market turbulence, according to Frank et al. (2022), is described as the inconstant and unpredictable changes in consumer preferences occurring within a company's markets, whereas technological turbulence refers to the rate of change in product technologies within a particular industry. In an environment characterized by significant regulatory turbulence, there exists a degree of uncertainty and unpredictability regarding laws and governmental regulations.

A turbulent environment imposes several limitations on a firm's current knowledge. Such environmental turbulence not only leads a firm to suffer the fate of competency traps because of responsiveness to current customer, but also disrupts its synergies among accumulated knowledge accompanied by organizational inertia for the new product. In turbulent environmental conditions, flexibly accommodating to environmental change and relentlessly renewing knowledge bases is the best way to sustain competitive advantage. In other words, environmental turbulence provides potential opportunities for firms to change or unlearn obsolete routines and competencies and scholars have also emphasized the need for organizations to change their existing belief structure, norms, and knowledge in a turbulent environment (Lyu et al., 2020). Due to the differences between technological innovativeness and market innovativeness, we argue that environmental turbulence has impact on business's performance when implementing digital marketing activities. Thus:

Proposition 6a: Environmental turbulence has moderating effect in the influence of digital marketing capabilities on digital marketing performance in SMEs.

Proposition 6b: Environmental turbulence has moderating effect in the influence of intellectual capital on digital marketing performance in SMEs.

The proposed conceptual model is shown in Figure 3 below.



3.3. Empirical questionnaire survey

3.3.1. Measurements scales

We further incorporate it with literature review, research measurements were developed and designed by referencing from other relevant previous studies. The measurements were adopted and modified to be suitable for the research. However, some items were newly developed based on the conceptual literature about this construct. To answer the research questions, we investigated factors that influence digital marketing performance, emphasizing the key role of digital marketing capabilities and intellectual capital. All items were rated on a 5-point Likert scale anchored at 1 = “strongly disagree” and 5 = “strongly agree.”

There are 9 items in *digital transformation* section which were adopted by Masrianto et al. (2022), Al Nuaimi et al. (2022) and Dabić et al. (2023). According to the scheme of this research, Wiklund and Shepherd (2003), *digital marketing capabilities* measure is comprised of 2 sub-constructs that contribute to digital marketing performance. The sub-construct of digital capabilities was adopted by Kamble et al. (2023) while those of marketing capabilities were adopted by the study of Mikalef et al. (2023) and Greenley et al. (2005), with 6 items for each sub-construct. On the *intellectual capital* scale, there were totally 23 items: 6 for human capital, 6 for structural capital, 6 for social capital and 5 for renewal capital, which emerged from Ahmed et al. (2022), Al-Omouh et al. (2022), Gerke et al. (2021), Gölgeci & Kuivalainen (2020), Kucharska (2022) and Kianto et al. (2010). Measurement items for *digital marketing performance* were adopted from Lamberti and Noci (2010) and Cao and Weerawardena (2023). A total group of 6

items were measured for the moderating variable scale: 3 items assessing *technological turbulence* and 3 items measuring *market turbulence*, adapted from Wang et al. (2022). The measurement scale is shown on Table 3.

Table 3. Measurement scale / Questionnaire Item			
Factor	No.	Item Descriptions	Reference
Digital Transformation (DT)	DT1	Our company's management has a vision to digitally transform for the future.	Masrianto et al. (2022)
	DT2	Our company aims to digitalize everything that can be digitalized.	Al Nuaimi et al. (2022)
	DT3	Our company aims to create more robust networking with digital technologies between the different business processes	Al Nuaimi et al. (2022)
	DT4	Our company aims to enhance an efficient customer interface and achieving information exchange with digitality.	Al Nuaimi et al. (2022)
	DT5	Our company promotes the cultural changes necessary for digital transformation.	Masrianto et al. (2022)
	DT6	Our company has launched a new business model based on/using digital technology.	Masrianto et al. (2022)
	DT7	Our company defines it by clear roles and responsibilities between departments to carry out digital initiatives.	Masrianto et al. (2022)
	DT8	Our company uses digital channels (such as online, social media, and mobile) to market its products and services.	Masrianto et al. (2022)
	DT9	Our company provide access for customers to engage into marketing strategies.	Dabić et al. (2023)
Digital Marketing Capabilities (DMC)	Digital Capabilities (DC)		
	DC1	Our company has interoperability in which information is exchanged and shared with the visibility and traceability of the processes.	Kamble et al. (2023)
	DC2	Our company has computing capability with reliable high-performance computing and data integration, storage and visualization.	Kamble et al. (2023)
	DC3	Our company has digitalized capability with improved working conditions for the organizational development.	Kamble et al. (2023)
	DC4	Our company has a certain degree of virtualization with simulation and modeling capabilities.	Kamble et al. (2023)
	DC5	Our company has cognitive with early fault detection and prediction, process monitoring, autonomous	Kamble et al. (2023)

		system, self-configurable, self-reasoning systems or self-predictive, self-sensing systems.	
	DC6	Our company has cyber security policies with data authentication and encryption or espionage, threat detection and monitoring or cyber security awareness programs.	Kamble et al. (2023)
	Marketing Capabilities (MC)		
	MC1	Our company has capabilities in gathering information about customers and competitors to track customer needs and analyze market information.	Mikalef et al. (2023)
	MC2	Our company has skills and ability to effectively segment and target market and develop creative marketing strategies	Mikalef et al. (2023)
	MC3	Our company has capabilities to allocate marketing resources effectively and translate and execute marketing strategies into action quickly.	Mikalef et al. (2023)
	MC4	Our company performs well in understanding what customer needs	Greenley et al. (2005)
	MC5	Our company performs well in creating, maintaining and enhancing relationships with key customers	Greenley et al. (2005)
	MC6	Our company performs well in achieving prices which attract customers and achieve financial goals	Greenley et al. (2005)
Intellectual Capital (IC)	Human Capital (HuC)		
	HuC1	Employees in our company hold suitable work experience for accomplishing their job successfully	Ahmed et al. (2022)
	HuC2	Employees in our company are specialists in their jobs and have suitable education to fulfill their jobs	Ahmed et al. (2022)
	HuC3	Employees in our company understand that doing this job well is a reward in itself.	Ahmed et al. (2022)
	HuC4	Considering the time spent on the job, employees in our company feel thoroughly familiar with their tasks.	Ahmed et al. (2022)
	HuC5	Employees in our company clearly recognize the company's visions and strategies.	Ahmed et al. (2022)
	HuC6	Employees in our company can effectively share their knowledge with each other.	Ahmed et al. (2022)
	Structural Capital (StC)		
	StC1	Our company realizes the relationships among authority, responsibility and benefit.	Ahmed et al. (2022)
	StC2	Our company effectively constructs and utilizes information system.	Ahmed et al. (2022)
	StC3	Our company has an efficient business process	Ahmed et al. (2022)
	StC4	Our company has active documented policies, instructions, standard procedures, and rules to support business operations	Al-Omouh et al. (2022)

	StC5	Much of our information and knowledge is embedded in the company's structure, manuals, archives, and databases.	Al-Omoush et al. (2022)
	StC6	Our company provides all the necessary tools, technologies, and facilities to support access to existing documents and information	Al-Omoush et al. (2022)
	Social Capital (SoC)		
	SoC1	Our company has long-term relationships with other partners.	Al-Omoush et al. (2022)
	SoC2	Our company shares the same ambitions and vision with other partners	Gerke et al. (2021)
	SoC3	Our company share consistent or compatible business goals with other partners	Gerke et al. (2021)
	SoC4	Our company encourages sharing knowledge and collaborating with others	Al-Omoush et al. (2022)
	SoC5	Our company has possibility to access and leverage valuable knowledge and resources through business relationships	Gölgeci & Kuivalainen (2020)
	SoC6	Our company has a good reputation for fair dealing	Gölgeci & Kuivalainen (2020)
	Renewal Capital (RnC)		
	RnC1	Our company has acquired a great deal of new and important knowledge.	Kucharska (2022)
	RnC2	Our employees have acquired many important skills and abilities.	Kucharska (2022)
	RnC3	Our company can be described as a learning organization.	Kucharska (2022)
	RnC4	The operations of our company can be described as creative and inventive.	Kucharska (2022)
	RnC5	Things that we have learnt have improved the performance of the organization	Kianto et al. (2010)
Digital Marketing Performance (DMP)	DMP1	Relative to our key competitors, our company has stronger growth in sales revenue from digital platforms with greater market share	Cao & Weerawardena (2023)
	DMP2	Relative to our key competitors, our company is more efficiently transform marketing inputs into marketing outputs in digital platforms.	Lamberti & Noci (2010)
	DMP3	Relative to our key competitors, our company is able to attract and retain profitable customers	Lamberti & Noci (2010)
	DMP4	Relative to our key competitors, our company is able to spread and maintain a strategic alignment with all of non-marketing functions within the firm as well as all the relevant supply-chain partners (channels, suppliers, outsourcers etc.)	Lamberti & Noci (2010)

	DMP5	Relative to our key competitors, our company is able to generate market intelligence, to spread it all over the organization and to spread a customer-oriented culture within the firm.	Lamberti & Noci (2010)
Environmental Turbulence (ET)	Technological Turbulence (TT)		
	TT1	The technology is changing rapidly	Wang et al., 2022)
	TT2	Technological changes provide big opportunities.	Wang et al., 2022)
	TT3	A large number of new products, services, or business models have been made possible through technological breakthroughs.	Wang et al., 2022)
	Market Turbulence (MT)		
	MT1	In this market, customers' preferences change quite a bit over time	Wang et al., 2022)
	MT2	Customers in this market are very receptive to new products, services, or business models ideas.	Wang et al., 2022)
	MT3	New customers tend to have products, services or business models- related needs that are different from those of existing customers.	Wang et al., 2022)

3.3.2. Data collection of empirical survey

As the present study focused on Vietnam, the survey approach was deemed appropriate for gathering the primary data since digital marketing is an emerging concept within Vietnamese businesses. As a result, we drew our sample from SMEs in Vietnam that are listed in the SMEs associations' websites. As our focus was on the SMEs with a good understanding of digital marketing-oriented business practices, we endeavored to choose respondent firms from all industries, whereby single respondents were selected from each company. High-level management (such as CEOs, directors, senior managers, etc.) were selected as respondents due to their direct involvement in managing organizational affairs and long-term strategic decision-making, as well as their first-hand familiarity with the business and operational procedures of organizations.

The initial survey questionnaire was developed after a thorough review of the literature and in-depth interviews and was provided to some senior-level managers and professors whose feedback was sought to ensure content validity. The survey instrument was prepared in both Vietnamese and English after a few minor adjustments. In order to mitigate the potential for various types of bias that are frequently observed in survey-based research, Spector and Brannick's (1995) recommendations were adhered to, whereby the independent and dependent variables were positioned far apart in the survey, the instrument was kept as brief as feasible, and great attention was paid to the wording. We randomly chose more than 700 businesses and emailed a copy of the questionnaire to each of their email addresses in order to get a sufficient sample size, along with a link to the survey where the completed form can be returned.

312 of responses were received; however, 16 responses were deemed inappropriate and were not included in the subsequent analysis. Consequently, 296 of firms, accounting for 94.87 percent of the responses, made up the final sample for data analysis.

4. Result of data analysis and findings

4.1. Reliability and validity

Table 4. Measurement properties

Construct	Identifier	Cronbach's Alpha	CR	AVE	Item	Factor loading
Digital Transformation	DT	0.927	0.943	0.735	DT1	0.677
					DT3	0.820
					DT5	0.807
					DT6	0.799
					DT8	0.864
					DT9	0.764
*Items DT2, DT4, DT7 were removed, the alpha-if-item-deleted method is applied						
Digital Marketing Capabilities	DMC	0.940	0.950	0.704	DC1	0.888
					DC2	0.576
					DC3	0.866
					DC5	0.756
					MC1	0.798
					MC2	0.566
					MC3	0.786
					MC5	0.735
*Items DC4 & MC4 were removed, the alpha-if-item-deleted method is applied						
Intellectual Capital	IC	0.963	0.967	0.663	HuC1	0.674
					HuC3	0.666
					HuC5	0.716
					HuC6	0.793
					StC2	0.806
					StC4	0.804
					StC6	0.860
					SoC1	0.765
					SoC2	0.881
					SoC4	0.871
					SoC5	0.860
					RnC1	0.869
					RnC2	0.832
					RnC3	0.831
RnC5	0.685					
*Items HuC2, HuC4, StC1, StC3, StC5, SoC3, SoC6 and RnC4 were removed, the alpha-if-item-deleted method is applied						
Digital Marketing Performance	DMP	0.941	0.958	0.850	DMP1	0.658
					DMP2	0.747
					DMP3	0.649

					DMP4	0.733
*Item DMP5 were removed, the alpha-if-item-deleted method is applied						
Environmental Turbulence	ET	0.951	0.951	0.803	TT1	0.749
					TT2	0.607
					TT3	0.705
					MT1	0.592
					MT2	0.622
					MT3	0.544
CR: Composite Reliability AVE: Average Variance Extracted						

We used Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) to examine the reliability. As shown in Table 4, all coefficients and CR were greater than 0.7, which indicated that all constructs had high internal consistency and reliability. Moreover, all AVE values were greater than 0.5 providing evidence of good convergent validity. In addition, AVE can also be used to evaluate discriminant validity. Overall, the examination of the scale's measurement properties shows conceptual consistency.

Table 5. Result of KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.930
Bartlett's Test of Sphericity	Approx. Chi-Square	19880.191
	df	741
	Sig.	0.000

In this study, the Kaiser-Meyer-Olkin (KMO) test is employed to assess the suitability of the data for factor analysis. This test evaluates the sampling adequacy for both individual variables and the overall model. The KMO statistic indicates the proportion of variance among variables that may be attributed to shared variance. A higher KMO value signifies a greater proportion of shared variance, indicating that the data is more appropriate for factor analysis. Result shows that KMO value is 0.930 indicating the data is very suitable for factor analysis (Kaiser, 1974).

Table 6. Latent covariance

	Mean	SD	DMC	DMP	DT	ET	IC
DMC	3.540	1.063	1	0.685	0.760	0.666	0.675
DMP	3.549	1.029	0.685	1	0.544	0.850	0.757
DT	3.633	1.079	0.760	0.544	1	0.574	0.582
ET	3.367	1.216	0.666	0.850	0.574	1	0.781
IC	3.685	1.081	0.675	0.757	0.582	0.781	1

Table 7. Results of Discriminant Validity (Fornell-Larcker Criterion)

	DMC	DMP	DT	ET	IC
DMC	0.839				
DMP	0.685	0.922			
DT	0.760	0.544	0.857		
ET	0.666	0.850	0.574	0.896	
IC	0.675	0.757	0.582	0.781	0.815

Table 8. Results of Discriminant Validity (Cross Loadings)

	DMC	DMP	DT	ET	IC
DC_01	0.794	0.426	0.677	0.377	0.416
DC_01	0.794	0.426	0.677	0.377	0.416
DC_02	0.806	0.558	0.826	0.573	0.566
DC_02	0.806	0.558	0.826	0.573	0.566
DC_03	0.822	0.471	0.647	0.429	0.466
DC_03	0.822	0.471	0.647	0.429	0.466
DC_05	0.855	0.545	0.582	0.503	0.496
DC_05	0.855	0.545	0.582	0.503	0.496
MC_01	0.873	0.565	0.538	0.551	0.555
MC_01	0.873	0.565	0.538	0.551	0.555
MC_02	0.849	0.716	0.709	0.698	0.702
MC_02	0.849	0.716	0.709	0.698	0.702
MC_03	0.864	0.641	0.567	0.634	0.621
MC_03	0.864	0.641	0.567	0.634	0.621
MC_05	0.845	0.642	0.578	0.657	0.67
MC_05	0.845	0.642	0.578	0.657	0.670
DMP_01	0.657	0.938	0.518	0.783	0.731
DMP_02	0.545	0.856	0.404	0.712	0.590
DMP_03	0.676	0.943	0.541	0.813	0.748
DMP_04	0.642	0.947	0.533	0.820	0.711
DT_01	0.694	0.623	0.83	0.644	0.691

DT_03	0.564	0.454	0.853	0.486	0.511
DT_05	0.680	0.506	0.94	0.525	0.519
DT_06	0.640	0.350	0.818	0.374	0.374
DT_08	0.678	0.342	0.797	0.372	0.380
DT_09	0.634	0.470	0.898	0.502	0.457
MT_01	0.620	0.694	0.528	0.908	0.695
MT_02	0.562	0.764	0.488	0.853	0.686
MT_03	0.635	0.667	0.525	0.882	0.693
TT_01	0.602	0.869	0.542	0.929	0.737
TT_02	0.610	0.766	0.499	0.877	0.675
TT_03	0.559	0.778	0.508	0.926	0.710
HuC_01	0.594	0.567	0.484	0.691	0.812
HuC_01	0.594	0.567	0.484	0.691	0.812
HuC_03	0.572	0.592	0.472	0.703	0.799
HuC_03	0.572	0.592	0.472	0.703	0.799
HuC_05	0.533	0.519	0.459	0.644	0.802
HuC_05	0.533	0.519	0.459	0.644	0.802
HuC_06	0.619	0.708	0.547	0.728	0.860
HuC_06	0.619	0.708	0.547	0.728	0.860
RnC_01	0.432	0.509	0.381	0.502	0.763
RnC_01	0.432	0.509	0.381	0.502	0.763
RnC_02	0.588	0.706	0.497	0.652	0.842
RnC_02	0.588	0.706	0.497	0.652	0.842
RnC_03	0.403	0.538	0.309	0.511	0.733
RnC_03	0.403	0.538	0.309	0.511	0.733
RnC_05	0.481	0.662	0.367	0.568	0.663
RnC_05	0.481	0.662	0.367	0.568	0.663
SoC_01	0.473	0.487	0.427	0.544	0.790
SoC_01	0.473	0.487	0.427	0.544	0.790
SoC_02	0.577	0.659	0.498	0.640	0.868
SoC_02	0.577	0.659	0.498	0.640	0.868
SoC_04	0.571	0.673	0.50	0.651	0.864
SoC_04	0.571	0.673	0.50	0.651	0.864

SoC_05	0.439	0.511	0.363	0.511	0.775
SoC_05	0.439	0.511	0.363	0.511	0.775
StC_02	0.637	0.692	0.588	0.717	0.876
StC_02	0.637	0.692	0.588	0.717	0.876
StC_04	0.642	0.687	0.591	0.725	0.888
StC_04	0.642	0.687	0.591	0.725	0.888
StC_06	0.608	0.673	0.530	0.683	0.851
StC_06	0.608	0.673	0.530	0.683	0.851

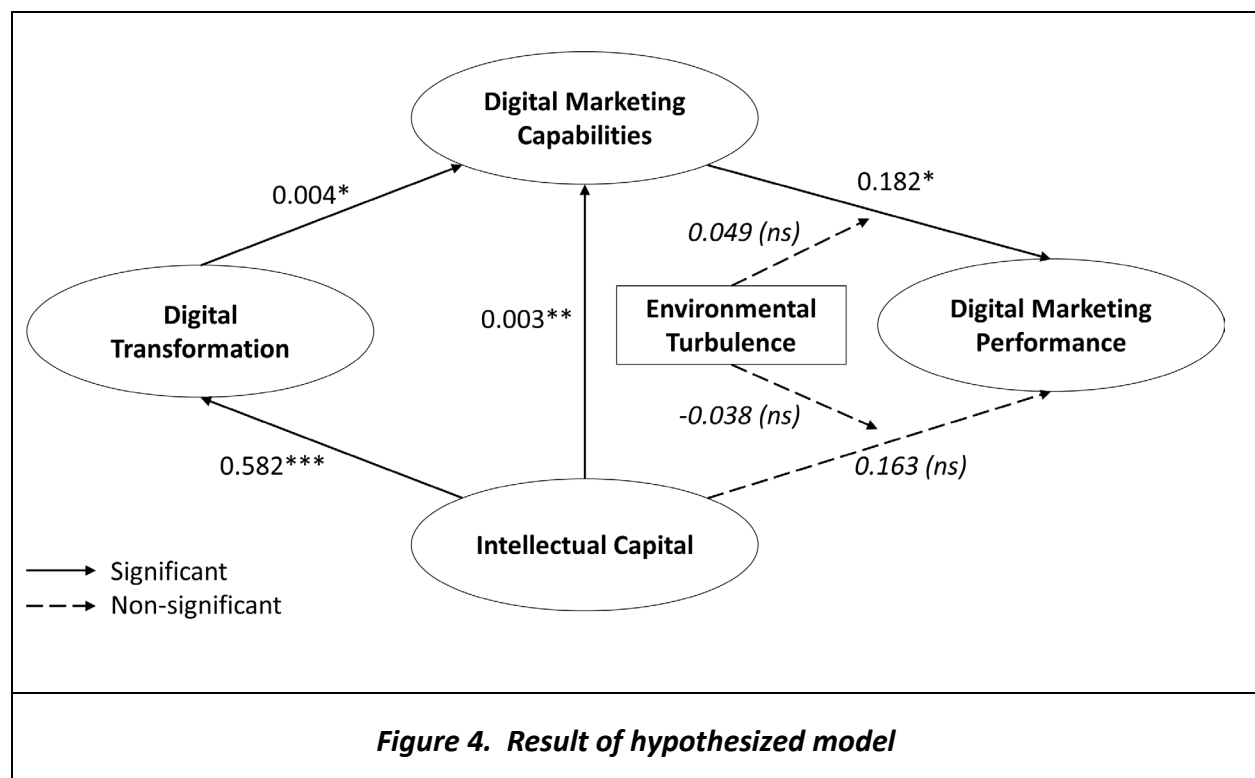
4.2. Regression models and hypotheses testing results

Table 9. Standardized path coefficients

Hypothesized Relationship		Path Coefficient	T Statistics	p-value	Result
DT → DMC	H1	0.004	2.184	0.029*	Supported
DMC → DMP	H2	0.182	2.291	0.022*	Supported
IC → DT	H3	0.582	11.882	0.000***	Supported
IC → DMC	H4	0.003	2.707	0.007**	Supported
IC → DMP	H5	0.163	1.715	0.087	Not supported
DMC * ET → DMP	H6a	0.049	0.850	0.395	Not supported
IC * ET → DMP	H6b	-0.038	0.808	0.419	Not supported
DC → DMC		0.506	46.037	0.000***	
MC → DMC		0.579	29.333	0.000***	
HuC → IC		0.295	26.315	0.000***	
RnC → IC		0.252	29.737	0.000***	
SoC → IC		0.269	38.628	0.000***	
StC → IC		0.284	21.285	0.000***	
Sample size:		N = 296			
R ²	DMC	1.000	Average R ²		0.7745
	DMP	0.760			
	DT	0.338			
	IC	1.000			
Average communality (AVE)		0.751	Goodness of Fit		0.7626
Note: a) *p < 0.05; **p < 0.01; ***p < 0.001.					
b) Goodness of fit (GoF) = $\sqrt{[(\text{Average communality}) \times (\text{Average } R^2)]}$.					
c) AVE, average variance extracted; DT, digital transformation; DMC, digital marketing capabilities; DC, digital capabilities; MC, marketing capabilities; IC, intellectual capital; HuC, human capital; StC, structural capital; SoC, social capital; RnC, renewal capital; DMP, digital marketing performance; TT, technological turbulence.					

SmartPLS was chosen as the most appropriate technique for analyzing our model. The results for direct effects are shown in Figure 3. As shown in Table 9, digital transformation ($\beta = 0.004$, $p < 0.05$) and intellectual capital ($\beta = 0.003$, $p < 0.01$) have positive significant influences on digital marketing capabilities, while digital marketing capabilities ($\beta = 0.182$, $p < 0.05$) have significant influence on digital marketing performance. Moreover, result also shows that intellectual capital ($\beta = 0.582$, $p < 0.001$) positively influenced digital transformation. Thus, H1, H2, H3 and H4 were supported. Intellectual capital were reported to have no direct significant influence on digital marketing performance ($p > 0.05$), which indicated that H5 were not supported. In addition, the result also shows that there were no significant moderating effects of environmental turbulence on the influence of digital marketing capabilities and intellectual on digital marketing

performance $p > 0.05$). Thus, H6a and H6b werenot supported. Additionally, this study also confirms the apparent sub-constructs of digital capability ($\beta = 0.506$, $p < 0.001$) and marketing capabililty ($\beta = 0.579$, $p < 0.001$) in contributing to the construct of digital marketing capabiilities; while intellectual capability is the combination of human capability ($\beta = 0.295$, $p < 0.001$), structural capability ($\beta = 0.284$, $p < 0.001$), social capability ($\beta = 0.269$, $p < 0.001$) and relational capability ($\beta = 0.252$, $p < 0.001$). An important part of model evaluation is examining fit indices that reflect the predictive power of estimated inner and outer model relationships. In line with the effect sizes for R^2 in the Goodness of fit (Wetzels, Odekerken-Schröder, and Van Oppen 2009), Table 9 shows that GoF = 0.7626, so the hypotheses were considered to be satisfied.



4.3. Theoretical contributions:

This research applies the Theory of the Resource-Based View (RBV), which is suited for analyzing the influence of the heterogeneousness of organizations on their performance as it aims to explain the competitiveness of individual firms by differences in their resources rather than market characteristics. The empirical testing was conducted to specifically determine how SMEs can utilize digital transformation into competitive advantages through enhancing digital marketing capabilities to have great impacts to the performance. The following important contributions to the literature are made.

Firstly, this research provides empirical results on digital marketing capabilities of SMEs that combines the constructs of digital and marketing capabilities of business. Support to the

argument of Hagen et al. (2022) indicating that studies investigating the factors that influence the uptake of digital marketing activities are still lacking, based on research results, this study contributes to the understanding of digital marketing capabilities of the existing literature on what digital marketing capabilities comprise and what the measurements and consequences of digital marketing capabilities. This study's findings help bridge the gaps identified by Quinton et al. (2018), supporting the idea that a key factor in a firm's success with digital marketing is the integration of digital capability and marketing capability. As shown by the current research, a firm's success in digital marketing hinges on the alignment of its digital and marketing capabilities. The digital capability offers the technological foundation (including tools, platforms, and analytics) necessary to facilitate data-driven, automated, and scalable processes. Effective marketing capability ensures that these tools are utilized effectively through the implementation of a well-planned strategy, innovative creativity, and a deep understanding of human insight. When these capabilities function in coordination, businesses can provide customized customer experiences, enhance operational productivity, and boost marketing return on investment. Organizations that successfully integrate their operations gain a substantial competitive edge within the constantly changing digital environment.

Secondly, this research makes a significant contribution to existing studies by examining the concept in a business-to-business small and medium-sized enterprise setting. The context may be familiar, but applying it from a new RBV viewpoint significantly boosts its relevance and credibility. Support to Shams et al. (2024) indicating that there are various digital marketing tools and capabilities to improve brand identity, which help in establishing close association with customers to influence their purchasing decisions in this digital era, current studies concentrate on the potential capabilities that companies can acquire through the effects of digital transformation and intellectual capital within the context of B2B in emerging economies, an area anticipated to undergo significant market changes over the next few years as a result of advancements in newly developed technologies. This study enriches understanding and creates a room to encourage researchers to delve into the subtleties of digital marketing capabilities and shed light on the factors influencing technology adoption and its effectiveness.

This study's findings also contribute to the understanding of how intellectual capital influences digital transformation, which in turn has implications for digital marketing capabilities, potentially leading to enhanced performance outcomes. The interplay relationship between intellectual capital, digital transformation, and digital marketing capabilities holds significant importance for organizations seeking to achieve exceptional results. Intellectual capital, consisting of human, structural, and relational components, forms the basis for facilitating digital transformation. This transformation, as a result, strengthens digital marketing capabilities, allowing companies to achieve more effective customer interaction, streamlined operational processes, and improved marketing outcomes. The path from intellectual capital or digital transformation to digital marketing performance, thus, is a continuous and a dynamic process and is not direct, in which intellectual capital facilitates digital transformation by equipping companies with the necessary technical infrastructure, skills, and processes to adapt and evolve; while digital transformation unleashes new digital marketing capabilities, including data-driven decision-making, predictive analytics, personalization, and omnichannel experiences.

Consequently, this study contributes to the literature by emphasize that digital marketing capabilities serve as a mediating variable that connects the impact of digital transformation on marketing performance. This means that while digital transformation provides the tools, infrastructure, and technological advancements, it is the firm's ability to leverage these capabilities (DMC) that ultimately drives performance.

4.4. Managerial implications:

This study provides some implications for practitioners. First, empirical evidence has provided insights into the connection between digital transformation and digital marketing capabilities, with a direct link confirmed, thereby establishing the relationship as a driving force behind strategic transformation, which enables businesses to gain a competitive advantage by upgrading their digital marketing capabilities. Our research demonstrates that digital transformation has a positive impact on digital marketing capabilities in small and medium-sized enterprises, which are more susceptible to stringent environmental expectations from their stakeholders and are better positioned to absorb the associated risks. In short, as a firm undergoes more digital transformation, it gains more capabilities to implement digital marketing strategies and practices effectively. Top management of firms like SMEs may launch new business model based on or using digital technology so that the firm is able to have capabilities in gathering information about customers and competitors to track customer needs and analyze market information as well as allocate marketing resources effectively and translate and execute marketing strategies into action quickly. Digital technologies also can help company perform well in understanding what customer needs, thus become increasingly effective in creating, maintaining and enhancing relationships with key customers.

The research findings clearly demonstrate that intellectual capital has a significant impact on the successful implementation of digital transformation and the enhancement of digital marketing capabilities within an organization. It is a need for managers to pay more intention in meliorating organization's intellectual capital in terms of human, structural, social and renewal capital. Every employee in the company should not only be a specialist in their job and have suitable education to fulfill their job but also clearly recognize the company's visions and strategies, in which they can effectively share their knowledge with each other. Moreover, the firms need to build effectively constructs and utilize information system with active documented policies, instructions, standard procedures, and rules to support business operations.

Furthermore, this study suggests that the company should develop long-term partnerships with other organizations, foster the exchange of knowledge and collaborate with others, and seek opportunities to tap into and utilize valuable knowledge and resources via its business relationships. These practices allow businesses to maintain their competitive edge, promote innovation, and improve their operational effectiveness. These initiatives should be a top priority for companies due to several key reasons. Collaboration with other organizations through long-term partnerships enables the exchange of ideas, skills, and expertise, thereby fueling innovation. For example: Technology firms can enter into collaborations with academic institutions and

research organizations to gain access to the latest discoveries, innovative technologies, and highly skilled researchers.

5. Discussion and Conclusions

This study examines the relationship between digital transformation and digital marketing capabilities as well as intellectual capital, and the moderating effects of some different factors. Our primitive findings may provide important theoretical and practical implications for utilizing and managing digital marketing capabilities in the digital transformation process from a B2B perspective. Theoretically, this study can expand understanding of the relationship between a digital transformation and digital marketing capabilities, and the influences of intellectual capital on digital transformation through utilizing digital marketing capabilities to enhance digital marketing performance, which enable firms to create new value as well as gain more competitive advantages in the marketshare. This study's findings is the consideration of value-added that result from B2B marketing operations. Moreover, the theoretical and empirical foundation of B2B marketing based on digital technology from process of digital transformation, is a significant value-addition to the existing literature on industrial marketing management. Moreover, our development of insights that center around the two variables – namely, digital marketing capabilities, which is a combinations of the context of digital capabilities and marketing capabilities; and intellectual capital - constitutes another essential contribution to the knowledge base of B2B marketing management. This is a crucial contribution to the field, as research focused on these micro and meso variables in the context of B2B digital marketing is notably scarce. Additionally, the research also try to examine the role of some control variables on digital marketing performance in various ways. In current study, we have proposed a research model and conducted an semi-structure experts interviews to receive findings. We also develop a meaurment scale based on the semi-structure interviews and literature review. An experical survey in Vietnam SME companies was also conducted for more insights.

5.1. The important role of digital transformation in the new era of marketing

Marketing is being transformed digitally, allowing for personalized experiences, data-driven information, and improved operational productivity. Empowering brands enables them to be more agile, innovative, and customer-oriented, ultimately leading to increased growth and a competitive edge in an ever-changing digital environment. This study suggests digital transformation plays an pivotal role in shaping and enhancing SMEs compatitive advantages by fostering the digital marketing capabilities. Digital transformation leverages advanced technologies to enhance customer experience, drive efficiency, and foster innovation. For example, digital transformation enables businesses to gather and analyze large volumes of customer data, allowing for personalized marketing campaigns tailored to individual preferences and behaviors. Tools like AI, machine learning, and predictive analytics help firms deliver targeted messages, product recommendations, and personalized content, leading to higher engagement and customer satisfaction. Access to real-time data and advanced analytics facilitates smarter, evidence-based marketing decisions. Digital transformation allows firms to quickly adapt to new trends, emerging platforms, and shifts in stakeholder's behavior, maintaining a competitive edge;

while the adoption of emerging technologies may enable immersive marketing experiences. Digital transformation also can break down geographical barriers, allowing firms to reach a global market. With e-commerce, social media, and digital advertising, companies can enter new markets with fewer physical constraints. Utilizing digital tools enable firms to adopt more sustainable marketing practices, reducing reliance on physical materials like print media.

5.2. The significant influence of intellectual capital on digital transformation and digital marketing capabilities.

Intellectual capital is reported to be an important part in facilitating digital transformation and boosting organizations' digital marketing capabilities. Organizations possess a broad range of knowledge, skills, experience, and intellectual assets that significantly affect their capacity to adapt, innovate, and compete in today's digital environment. Moreover, intellectual capital plays a crucial role in both facilitating digital transformation and enhancing digital marketing capabilities. The skills and adaptability needed for change are powered by human capital, while structural capital establishes the framework for automation and data-driven marketing. Relational capital enhances stakeholder engagement, and knowledge capital fuels innovation and ongoing learning. Collectively, these components enable organizations to boost marketing effectiveness, improve customer satisfaction, and sustain a competitive advantage in the digital era. Firms that invest in intellectual capital are more favorably situated to tap into the full potential of digital transformation and attain long-term growth.

5.3. The median role of digital marketing capabilities in fostering digital marketing performance by digital transformation and intellectual capital

Digital marketing capabilities act as a key intermediary in the connection between digital transformation, intellectual capital, and digital marketing performance. By connecting the strategic use of digital transformation to the efficient use of intellectual capital, DMCs allow companies to attain better marketing results. The integration of cutting-edge technologies into business operations results in changes to how things are done and generates value. Unless companies possess robust digital marketing capabilities, the full potential of digital transformation cannot be achieved. Digital transformation fosters a flexible business environment in which marketing teams are able to swiftly adapt to shifts in consumer habits, rising digital platforms, and market developments. Moreover, a company's intellectual capital consists of the non-visible knowledge, abilities, procedures, and connections that contribute to its market superiority. Companies can, therefore, build and maintain robust digital marketing capabilities by utilizing information communication technologies. For example: marketing teams are bolstered by the expertise of skilled professionals who have a grasp of SEO, data analysis, and digital platforms which can strengthen core marketing capabilities. Strong partnerships with customers, partners, and influencers enable businesses to collaborate on content development and leverage significant professional networks.

5.4. Limitation and future research

This study also remains some limitations. First, as the model and the hypotheses proposed have been tested on a sample composed only of Vietnamese companies, cross-nations design research can be conducted for a more comprehensive insight. Second, our study's limitations are partially due to the scope of the literature examined might be limited, potentially not covering all the latest developments in digital transformation applications for digital marketing. Like any literature review, potential biases in choosing sources and interpreting them may impact the results. Future research would be also interesting to bring off a design on longitudinal research to fully explore topics which are related to digital transformation and digital marketing capabilities that have impacts on many aspects of business operations.

Additionally, future investigations should concentrate on examining the implementation of digital transformation in marketing and its effects on marketing professionals. Developing a cohesive framework is essential for the structured execution of digital transformation in marketing, linking relevant digital tools to corresponding marketing areas and assessing the extent of their interconnectedness. It is essential to include examples of digital technologies that target various marketing objectives in order to demonstrate how digital transformation can go beyond conventional frameworks and lead to new possibilities for research and practical implementation. Last but not least, future research may focus on the interplay between digital transformation, intellectual capital, and digital marketing capabilities, looking deeper to explore how different aspects of intellectual capital (human, structural, and relational) impacts digital marketing performance.

References

- Adams, P., Freitas, I. M. B., & Fontana, R. (2019). Strategic orientation, innovation performance and the moderating influence of marketing management. *Journal of Business Research*, 97, 129-140.
- Agostini, L., Nosella, A., & Filippini, R. (2017). Does intellectual capital allow improving innovation performance? A quantitative analysis in the SME context. *Journal of Intellectual Capital*, 18(2), 400-418.
- Ahmed, A., Bhatti, S. H., Gölgeci, I., & Arslan, A. (2022). Digital platform capability and organizational agility of emerging market manufacturing SMEs: The mediating role of intellectual capital and the moderating role of environmental dynamism. *Technological Forecasting and Social Change*, 177, 121513. <https://doi.org/10.1016/j.techfore.2022.121513>
- Al Nuaimi, B. K., Singh, S. K., Ren, S., Budhwar, P., & Vorobyev, D. (2022). Mastering digital transformation: The nexus between leadership, agility, and digital strategy. *Journal of Business Research*, 145, 636-648. <https://doi.org/10.1016/j.jbusres.2022.03.038>
- Al Qershi, N., Saufi, R. B. A., Ismail, N. A., Mohamad, M. R. B., Ramayah, T., Muhammad, N. M. N., & Yusoff, M. N. H. B. (2023). The moderating role of market turbulence beyond the Covid-19 pandemic and Russia-Ukraine crisis on the relationship between intellectual

- capital and business sustainability. *Technological Forecasting and Social Change*, 186, 122081. <https://doi.org/10.1016/j.techfore.2022.122081>
- Al-Omoush, K. S., Palacios-Marques, D., & Ulrich, K. (2022). The impact of intellectual capital on supply chain agility and collaborative knowledge creation in responding to unprecedented pandemic crises. *Technological Forecasting and Social Change*, 178, 121603. <https://doi.org/10.1016/j.techfore.2022.121603>
- Andonova, V., & Losada-Otalora, M. (2020). Understanding the interplay between brand and innovation orientation: Evidence from emerging multinationals. *Journal of Business Research*, 119, 540-552.
- Annarelli, A., Battistella, C., Nonino, F., Parida, V., & Pessot, E. (2021). Literature review on digitalization capabilities: Co-citation analysis of antecedents, conceptualization and consequences. *Technological Forecasting and Social Change*, 166, 120635.
- Arshad, M. Z., Arshad, D., Lamsali, H., Alshuaibi, A. S. I., Alshuaibi, M. S. I., Albashar, G., ... & Chuah, L. F. (2023). Strategic resources alignment for sustainability: The impact of innovation capability and intellectual capital on SME's performance. Moderating role of external environment. *Journal of Cleaner Production*, 417, 137884. <https://doi.org/10.1016/j.jclepro.2023.137884>
- Bachmann, J. T., Ohlles, I., & Flatten, T. (2021). Effects of entrepreneurial marketing on new ventures' exploitative and exploratory innovation: The moderating role of competitive intensity and firm size. *Industrial Marketing Management*, 92, 87-100. <https://doi.org/10.1016/j.indmarman.2020.10.002>
- Biemans, W. (2023). The impact of digital tools on sales-marketing interactions and perceptions. *Industrial Marketing Management*, 115, 395-407. <https://doi.org/10.1016/j.indmarman.2023.10.015>
- Behl, A., Jayawardena, N., Nigam, A., Pereira, V., Shankar, A., & Jebarajakirthy, C. (2023). Investigating the revised international marketing strategies during COVID-19 based on resources and capabilities of the firms: A mixed method approach. *Journal of Business Research*, 158, 113662.
- Cao, G., & Tian, N. (2020). Enhancing customer-linking marketing capabilities using marketing analytics. *Journal of Business & Industrial Marketing*, 35(7), 1289-1299.
- Cao, G., & Weerawardena, J. (2023). Strategic use of social media in marketing and financial performance: The B2B SME context. *Industrial Marketing Management*, 111, 41-54. <https://doi.org/10.1016/j.indmarman.2023.03.007>
- Chen, K., & Zhang, S. (2024). How does open public data impact enterprise digital transformation? *Economic Analysis and Policy*. Volume 83, September 2024, Pages 178-190.
- Chou, C., & Yang, K. P. (2011). The interaction effect of strategic orientations on new product performance in the high-tech industry: A nonlinear model. *Technological Forecasting and Social Change*, 78(1), 63-74. <https://doi.org/10.1016/j.techfore.2010.07.018>
- Chu, Y., Chi, M., Wang, W., & Luo, B. (2019). The impact of information technology capabilities of manufacturing enterprises on innovation performance: Evidences from SEM and fsQCA. *Sustainability*, 11(21), 5946.
- Collis, D. J., & Montgomery, C. A. (2009). Competing on Resources: Strategy in the 1990s. In *Knowledge and strategy* (pp. 25-40). Routledge.

- Dabić, M., Posinković, T. O., Vlačić, B., & Gonçalves, R. (2023). A configurational approach to new product development performance: The role of open innovation, digital transformation and absorptive capacity. *Technological Forecasting and Social Change*, 194, 122720. <https://doi.org/10.1016/j.techfore.2023.122720>
- Day, G. S. (1994). The capabilities of market-driven organizations. *Journal of Marketing*, 58(4), 37-52.
- Degbey, W. Y., & Pelto, E. (2021). Customer knowledge sharing in cross-border mergers and acquisitions: The role of customer motivation and promise management. *Journal of International Management*, 27(4), 100858.
- Deng, C., Yang, J., Su, Z., & Zhang, S. (2021). The double-edged sword impact of effectuation on new product creativity: The moderating role of competitive intensity and firm size. *Journal of Business Research*, 137, 1-12.
- Deng, G., Zhang, J., & Xu, Y. (2024). How e-commerce platforms build channel power: the role of AI resources and market-based assets. *Journal of Business & Industrial Marketing*, 39(2), 173-188.
- Dethine, B., Enjolras, M., & Monticolo, D. (2020). Digitalization and SMEs' export management: Impacts on resources and capabilities. *Technology Innovation Management Review*, 10(4). <http://doi.org/10.22215/timreview/1344>
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic Management Journal*, 21(10-11), 1105-1121.
- Elia, S., Giuffrida, M., Mariani, M. M., & Bresciani, S. (2021). Resources and digital export: An RBV perspective on the role of digital technologies and capabilities in cross-border e-commerce. *Journal of Business Research*, 132, 158-169.
- EMG (2023, August 24). *Decode, drive, deliver: Mastering digital marketing performance analysis for optimal success*. LinkedIn. Retrieved on July 27, 2024 at <https://www.linkedin.com/pulse/decode-drive-deliver-mastering-digital-marketing-performance>
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.
- Farzaneh, M., Wilden, R., Afshari, L., & Mehralian, G. (2022). Dynamic capabilities and innovation ambidexterity: The roles of intellectual capital and innovation orientation. *Journal of Business Research*, 148, 47-59. <https://doi.org/10.1016/j.jbusres.2022.04.030>
- Feliciano-Cestero, M. M., Ameen, N., Kotabe, M., Paul, J., & Signoret, M. (2023). Is digital transformation threatened? A systematic literature review of the factors influencing firms' digital transformation and internationalization. *Journal of Business Research*, 157, 113546. <https://doi.org/10.1016/j.jbusres.2022.113546>
- Ferreira, J. J., Fernandes, C. I., & Ferreira, F. A. (2019). To be or not to be digital, that is the question: Firm innovation and performance. *Journal of Business research*, 101, 583-590. <https://doi.org/10.1016/j.jbusres.2018.11.013>
- Frank, A. G., de Souza Mendes, G. H., Benitez, G. B., & Ayala, N. F. (2022). Service customization in turbulent environments: Service business models and knowledge integration to create capability-based switching costs. *Industrial Marketing Management*, 100, 1-18.

- Garzoni, A., De Turi, I., Secundo, G., & Del Vecchio, P. (2020). Fostering digital transformation of SMEs: a four levels approach. *Management Decision*, 58(8), 1543-1562.
<https://doi.org/10.1108/MD-07-2019-0939>
- Gerke, A., Luzzini, D., & Mena, C. (2021). Innovation configurations in sport clusters: The role of interorganizational citizenship and social capital. *Journal of Business Research*, 133, 409-419. <https://doi.org/10.1016/j.jbusres.2021.04.064>
- Gölgeci, I., & Kuivalainen, O. (2020). Does social capital matter for supply chain resilience? The role of absorptive capacity and marketing-supply chain management alignment. *Industrial Marketing Management*, 84, 63-74. <https://doi.org/10.1016/j.indmarman.2019.05.006>
- Greenley, G. E., Hooley, G. J., & Rudd, J. M. (2005). Market orientation in a multiple stakeholder orientation context: implications for marketing capabilities and assets. *Journal of Business Research*, 58(11), 1483-1494. <https://doi.org/10.1016/j.jbusres.2004.07.004>
- Gregory, G. D., Ngo, L. V., & Karavdic, M. (2019). Developing e-commerce marketing capabilities and efficiencies for enhanced performance in business-to-business export ventures. *Industrial Marketing Management*, 78, 146-157.
<https://doi.org/10.1016/j.indmarman.2017.03.002>
- Guo, H., Xu, H., Tang, C., Liu-Thompkins, Y., Guo, Z., & Dong, B. (2018). Comparing the impact of different marketing capabilities: Empirical evidence from B2B firms in China. *Journal of Business Research*, 93, 79-89. <https://doi.org/10.1016/j.jbusres.2018.04.010>
- Hagen, D., Risselada, A., Spierings, B., Weltevreden, J. W. J., & Atzema, O. (2022). Digital marketing activities by Dutch place management partnerships: A resource-based view. *Cities*, 123, 103548.
- Hashim, N., Taleb, T. S., & Faudzi, M. S. M. (2023). Entrepreneurial bricolage and micro-business performance: A moderated mediation of innovation capability and digital marketing capabilities. *Journal of Strategic Marketing*, 1-16.
- Haverila, M. J., & Haverila, K. C. (2024). The influence of quality of big data marketing analytics on marketing capabilities: the impact of perceived market performance!. *Marketing Intelligence & Planning*, 42(2), 346-372.
- Helfat, C. E., & Lieberman, M. B. (2002). The birth of capabilities: market entry and the importance of pre-history. *Industrial and Corporate Change*, 11(4), 725-760.
- Herhausen, D., Miočević, D., Morgan, R. E., & Kleijnen, M. H. (2020). The digital marketing capabilities gap. *Industrial Marketing Management*, 90, 276-290.
<https://doi.org/10.1016/j.indmarman.2020.07.022>
- Homburg, C., & Wielgos, D. M. (2022). The value relevance of digital marketing capabilities to firm performance. *Journal of the Academy of Marketing Science*, 50(4), 666-688.
- Hung, K. P., & Chou, C. (2013). The impact of open innovation on firm performance: The moderating effects of internal R&D and environmental turbulence. *Technovation*, 33(10-11), 368-380. <https://doi.org/10.1016/j.technovation.2013.06.006>
- Iankova, S., Davies, I., Archer-Brown, C., Marder, B., & Yau, A. (2019). A comparison of social media marketing between B2B, B2C and mixed business models. *Industrial Marketing Management*, 81, 169-179. <https://doi.org/10.1016/j.indmarman.2018.01.001>
- Järvinen, J. (2016). *The use of digital analytics for measuring and optimizing digital marketing performance* (Doctoral dissertation, University of Jyväskylä).

- Järvinen, J., & Karjaluoto, H. (2015). The use of Web analytics for digital marketing performance measurement. *Industrial Marketing Management*, 50, 117-127.
<https://doi.org/10.1016/j.indmarman.2015.04.009>
- Jaworski, B. J., & Lurie, R. S. (2019). Building marketing capabilities: Principles from the field. *AMS Review*, 9(3), 372-380. **<https://doi.org/10.1007/s13162-019-00160-w>**
- Kalubanga, M., & Gudergan, S. (2022). The impact of dynamic capabilities in disrupted supply chains—The role of turbulence and dependence. *Industrial Marketing Management*, 103, 154-169.
- Kamble, S., Rana, N. P., Gupta, S., Belhadi, A., Sharma, R., & Kulkarni, P. (2023). An effectuation and causation perspective on the role of design thinking practices and digital capabilities in platform-based ventures. *Technological Forecasting and Social Change*, 193, 122646.
<https://doi.org/10.1016/j.techfore.2023.122646>
- Kannan, P. K. & Li, H. (2017). Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing*, 34(1), 22-45.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Khattak, M. S., Wu, Q., Ahmad, M., & Anwar, M. (2024). Probing the impact of intellectual capital on SME efficiency: the role of business model innovation. *Journal of Intellectual Capital*. Vol. 25 No. 2/3, 2024 pp. 613-639.
- Kianto, A., Hurmelinna-Laukkanen, P., & Ritala, P. (2010). Intellectual capital in service- and product-oriented companies. *Journal of intellectual capital*, 11(3), 305-325.
- Kozlenkova, I. V., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. *Journal of The Academy of Marketing Science*, 42, 1-21.
- Krings, W., Palmer, R., & Inversini, A. (2021). Industrial marketing management digital media optimization for B2B marketing. *Industrial Marketing Management*, 93, 174-186.
<https://doi.org/10.1016/j.indmarman.2021.01.002>
- Kucharska, W. (2022). Tacit knowledge influence on intellectual capital and innovativeness in the healthcare sector: A cross-country study of Poland and the US. *Journal of Business Research*, 149, 869-883.
- Kumar, B., Sharma, A., Vatavwala, S., & Kumar, P. (2020). Digital mediation in business-to-business marketing: A bibliometric analysis. *Industrial Marketing Management*, 85, 126-140. **<https://doi.org/10.1016/j.indmarman.2019.10.002>**
- Lamberti, L., & Noci, G. (2010). Marketing strategy and marketing performance measurement system: Exploring the relationship. *European Management Journal*, 28(2), 139-152.
<https://doi.org/10.1016/j.emj.2009.04.007>
- Leeflang, P. S., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European management journal*, 32(1), 1-12.
- Li, H., Wu, Y., Cao, D., & Wang, Y. (2021). Organizational mindfulness towards digital transformation as a prerequisite of information processing capability to achieve market agility. *Journal of Business Research*, 122, 700-712.
<https://doi.org/10.1016/j.jbusres.2019.10.036>
- Lyu, C., Yang, J., Zhang, F., Teo, T. S., & Guo, W. (2020). Antecedents and consequence of organizational unlearning: Evidence from China. *Industrial Marketing Management*, 84, 261-270. **<https://doi.org/10.1016/j.indmarman.2019.07.013>**

- Madhavaram, S., Appan, R., Manis, K. T., & Browne, G. J. (2022). Building Capabilities for Software Development Firm Competitiveness: The Role of Intellectual Capital and Intra-Firm Relational Capital. *Information & Management*, 103744.
<https://doi.org/10.1016/j.im.2022.103744>
- Marsh, E., Vallejos, E. P., & Spence, A. (2022). The digital workplace and its dark side: An integrative review. *Computers in Human Behavior*, 128, 107118.
- Masrianto, A., Hartoyo, H., Hubeis, A. V. S., & Hasanah, N. (2022). Digital Marketing Utilization Index for Evaluating and Improving Company Digital Marketing Capability. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(3), 153.
<https://doi.org/10.3390/joitmc8030153>
- Mathew, V., & Soliman, M. (2021). Does digital content marketing affect tourism consumer behavior? An extension of the technology acceptance model. *Journal of Consumer Behaviour*, 20(1), 61-75.
- Mazzucchelli, A., Gurioli, M., Graziano, D., Quacquarelli, B., & Aouina-Mejri, C. (2021). How to fight against food waste in the digital era: Key factors for a successful food sharing platform. *Journal of Business Research*, 124, 47-58.
- McDowell, W. C., Peake, W. O., Coder, L., & Harris, M. L. (2018). Building small firm performance through intellectual capital development: Exploring innovation as the “black box”. *Journal of business research*, 88, 321-327.
- Meijerink, J., & Bondarouk, T. (2018). Uncovering configurations of HRM service provider intellectual capital and worker human capital for creating high HRM service value using fsQCA. *Journal of business research*, 82, 31-45.
- Melović, B., Jocović, M., Dabić, M., Vulić, T. B., & Dudic, B. (2020). The impact of digital transformation and digital marketing on the brand promotion, positioning and electronic business in Montenegro. *Technology in Society*, 63, 101425.
- Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, 164, 113998. **<https://doi.org/10.1016/j.jbusres.2023.113998>**
- Morgan, N. A., Clark, B. H., & Gooner, R. (2002). Marketing productivity, marketing audits, and systems for marketing performance assessment: integrating multiple perspectives. *Journal of business research*, 55(5), 363-375.
- Nordin, F., & Lindbergh, J. (2019). Foreign market learning: an integrative model of its antecedents, processes and outcomes. *Journal of Business & Industrial Marketing*, 34(6), 1248-1258.
- Nylén, D., & Holmström, J. (2015). Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation. *Business Horizons*, 58(1), 57-67.
<https://doi.org/10.1016/j.bushor.2014.09.001>
- Obradović, T., Vlačić, B., & Dabić, M. (2021). Open innovation in the manufacturing industry: A review and research agenda. *Technovation*, 102, 102221.
<https://doi.org/10.1016/j.technovation.2021.102221>
- Olson, E. M., Olson, K. M., Czaplewski, A. J., & Key, T. M. (2021). Business strategy and the management of digital marketing. *Business Horizons*, 64(2), 285-293.

- Pandey, N., Nayal, P., & Rathore, A. S. (2020). Digital marketing for B2B organizations: structured literature review and future research directions. *Journal of Business & Industrial Marketing*, 35(7), 1191-1204.
- Paolone, F., Cucari, N., Wu, J., & Tiscini, R. (2022). How do ESG pillars impact firms' marketing performance? A configurational analysis in the pharmaceutical sector. *Journal of Business & Industrial Marketing*, 37(8), 1594-1606.
- Prahalad, C. K. & Hamel G. (1990). *The Core Competence of the Corporation*. Harvard Business Review, 68, pp. 79-91. https://doi.org/10.1007/3-540-30763-X_14
- Rachinger, M., Rauter, R., Müller, C., Vorraber, W., & Schirgi, E. (2018). Digitalization and its influence on business model innovation. *Journal of Manufacturing Technology Management*, 30(8), 1143-1160. <https://doi.org/10.1108/JMTM-01-2018-0020>
- Ravat, L., Hemonnet-Goujot, A., & Hollet-Haudebert, S. (2024). Exploring how to develop data-driven innovation capability of marketing within B2B firms: Toward a capability model and process-oriented approach. *Industrial Marketing Management*, 118, 110-125.
- Rayburn, S. W., Badrinarayanan, V., Anderson, S. T., & Gupta, A. (2021). Continuous techno-training and business-to-business salesperson success: How boosting techno-efficacy enhances sales effort and performance. *Journal of Business Research*, 133, 66-78.
- Ritala, P., Kianto, A., Vanhala, M., & Hussinki, H. (2023). To protect or not to protect? Renewal capital, knowledge protection and innovation performance. *Journal of Knowledge Management*, 27(11), 1-24.
- Rizvanović, B., Zutshi, A., Grilo, A., & Nodehi, T. (2023). Linking the potentials of extended digital marketing impact and start-up growth: Developing a macro-dynamic framework of start-up growth drivers supported by digital marketing. *Technological Forecasting and Social Change*, 186, 122128. <https://doi.org/10.1016/j.techfore.2022.122128>
- Sakas, D. P., Giannakopoulos, N. T., Nasiopoulos, D. K., Kanellos, N., & Tsoulfas, G. T. (2023). Assessing the Efficacy of Cryptocurrency Applications' Affiliate Marketing Process on Supply Chain Firms' Website Visibility. *Sustainability*, 15(9), 7326.
- Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021). Setting B2B digital marketing in artificial intelligence-based CRMs: A review and directions for future research. *Industrial Marketing Management*, 98, 161-178.
- Setkute, J., & Dibb, S. (2022). "Old boys' club": Barriers to digital marketing in small B2B firms. *Industrial Marketing Management*, 102, 266-279. <https://doi.org/10.1016/j.indmarman.2022.01.022>
- Shams, R., Chatterjee, S., & Chaudhuri, R. (2024). Developing brand identity and sales strategy in the digital era: Moderating role of consumer belief in brand. *Journal of Business Research*, 179, 114689.
- Shukla, A., Varshney, J., & Raj, A. (2024). Examining the linkage between managerial ties and firm performance: The mediating role of marketing capabilities and moderation role of industry-A meta-analytic approach. *Industrial Marketing Management*, 119, 122-134.
- Spector, P. E., & Brannick, M. T. (1995). "The Nature and Effects of Method Variance in Organizational Research". In C. L. Cooper & I. T. Robinson (Eds.), *International Review of Industrial and Organizational Psychology*, John Wiley & Sons, p. 249-274. Available at https://digitalcommons.usf.edu/psy_facpub/568

- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509-533.
- Ulaga, W., & Reinartz, W. J. (2011). Hybrid offerings: How manufacturing firms combine goods and services successfully. *Journal of marketing*, 75(6), 5-23.
<https://doi.org/10.1509/jm.09.0395>
- Varadarajan, R. (2024). Resource-advantage theory, resource-based theory and market-based resources advantage: effect of marketing performance on customer information assets stock and information analysis capabilities. *Journal of Marketing Management*, 1-20.
- Verhoef, P. C., & Bijmolt, T. H. (2019). Marketing perspectives on digital business models: A framework and overview of the special issue. *International Journal of Research in Marketing*, 36(3), 341-349. <https://doi.org/10.1016/j.ijresmar.2019.08.001>
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 28(2), 118-144.
<https://doi.org/10.1016/j.jsis.2019.01.003>
- Vu, M. C., & Nguyen, L. A. (2022). Mindful unlearning in unprecedented times: Implications for management and organizations. *Management Learning*, 53(5), 797-817.
- Wang, C., Qureshi, I., Guo, F., & Zhang, Q. (2022). Corporate Social Responsibility and Disruptive Innovation: The moderating effects of environmental turbulence. *Journal of Business Research*, 139, 1435-1450. <https://doi.org/10.1016/j.jbusres.2021.10.046>
- Wang, F. (2020). Digital marketing capabilities in international firms: a relational perspective. *International Marketing Review*, 37(3), 559-577.
- Westerman, G., Bonnet, D., & McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Harvard Business Press.
<https://hbsp.harvard.edu/product/17039-HBK-ENG>
- Wilden, R., & Gudergan, S. P. (2015). The impact of dynamic capabilities on operational marketing and technological capabilities: investigating the role of environmental turbulence. *Journal of the academy of marketing science*, 43, 181-199.
- Witschel, D., Baumann, D., & Voigt, K. I. (2022). How manufacturing firms navigate through stormy waters of digitalization: The role of dynamic capabilities, organizational factors and environmental turbulence for business model innovation. *Journal of Management & Organization*, 28(3), 681-714.
- Wong, A. H. K., Wu, C. X., Whitla, P., & Snell, R. S. (2022). How LMX and marketing capabilities guide and motivate customer-facing employees' learning. *Journal of Business Research*, 138, 161-169.
- Wu, C. W., Botella-Carrubi, D., & Blanco-González-Tejero, C. (2024). The empirical study of digital marketing strategy and performance in small and medium-sized enterprises (SMEs). *Technological Forecasting and Social Change*, 200, 123142.
- Wu, L., Sun, L., Chang, Q., Zhang, D., & Qi, P. (2022). How do digitalization capabilities enable open innovation in manufacturing enterprises? A multiple case study based on resource integration perspective. *Technological Forecasting and Social Change*, 184, 122019.
<https://doi.org/10.1016/j.techfore.2022.122019>

- Yang, F., Luo, C., & Pan, L. (2024). Do digitalization and intellectual capital drive sustainable open innovation of natural resources sector? Evidence from China. *Resources Policy*, 88, 104345. [**https://doi.org/10.1016/j.resourpol.2023.104345**](https://doi.org/10.1016/j.resourpol.2023.104345)
- Yuk, H., & Garrett, T. C. (2023). Does customer participation moderate the effects of innovation on cost-based financial performance? An examination of different forms of customer participation. *Journal of Business Research*, 156, 113479. [**https://doi.org/10.1016/j.jbusres.2022.113479**](https://doi.org/10.1016/j.jbusres.2022.113479)
- Zeithaml, V. A., Jaworski, B. J., Kohli, A. K., Tuli, K. R., Ulaga, W., & Zaltman, G. (2020). A theories-in-use approach to building marketing theory. *Journal of Marketing*, 84(1), 32-51. [**https://doi.org/10.1177/0022242919888477**](https://doi.org/10.1177/0022242919888477)
- Zhang, F., & Zhu, L. (2021). Social media strategic capability, organizational unlearning, and disruptive innovation of SMEs: The moderating roles of TMT heterogeneity and environmental dynamism. *Journal of Business Research*, 133, 183-193.

The Three Pillars of Digital Transformation: Technology Integration, Platforms, and Skills

Ganesh Bhatt

Morgan State University

&

Samuel Ejiaku

Morgan State University

Abstract:

Since the mid-1990s, the academic community has increasingly focused on the multifaceted dimensions of digital transformation. This has led to a substantial rise in research publications about digital transformation and digital competences. Despite significant advancements, a coherent theory for conceptualizing elements of digital transformation and specifying diverse attributes of digital capabilities remains undeveloped. Digital transformation is not merely a technological move; it entails a comprehensive reconfiguration of organizational processes, capabilities, and culture. This paper analyzes the complex interrelationships among three essential components—technology integration, platform capabilities, and employee skills—in establishing a strong digital ecosystem that supports sustained digital transformation. Paper, also provides an overview of dynamic capabilities and digital transformation.

The Three Pillars of Digital Transformation: Technology Integration, Platforms, and Skills

Introduction:

Digital transformation is a prevalent term combining various concepts. However, its meanings and notions vary across different contexts. For some, digital transformation entails the adoption of cutting-edge technology or transitioning to cloud services (Kohli and Melville, 2019); for others, it signifies a fundamental reconfiguration of organizational operations (Burrus and Mann, 2011); and for yet others, it focuses on improving the customer experience. Therefore, it is asserted that firms should leverage technology to enhance customer relations by optimizing processes, automating jobs, and reducing costs (Gil-Gomez et al., 2020). Each author has presented a streamlined interpretation of digital transformation utilizing basic frameworks. The absence of a cohesive theory has significantly hindered the comprehension of digital transformation, as the study has been disjointed and lacks a cumulative tradition of building upon previous studies. The primary objective of this study is to shed light on the concept of digital transformation. An understanding of the digital competencies that facilitate digital transformation will benefit academic scholars and practicing managers.

Literature Review:

Multiple notions of digital transformation have emerged. A body of studies perceives digital transformation as a mechanism to foster corporate innovation and address client requirements (Hinterhuber and Liozu, 2021; Seymen, 2022; Shrivastava, 2017). An additional line of inquiry posits that digital transformation equates to value creation for enterprises (Cennamo et al., 2020;

Orji, 2019). Pratt and Boulton (2023) in CIO, a distinguished information systems magazine, offer a comparable definition, stating that digital transformation entails a major alteration in an organization's value delivery to its consumers. Kriess and Brennen (2016) characterize digital transformation as the democratization of information for the populace. Many of the above definitions assert that digital transformation facilitates innovative approaches, enhances efficiency in addressing conventional activities, and improves customer value delivery. Digital revolution renders traditional processes and communication outmoded, hence affecting market competitiveness (Schwertner, 2017). In "Roadmap for Digital Transformation," Zaoui and Souissi (2020) examine the worldwide importance of digital transformation for enterprises in every industry. Although digital transformation is appealing to businesses, a comprehensive perspective on the subject is still missing.

Vial (2021) contends that it is essential to meticulously analyze the characteristics and implications of digital transformation. The author analyzes 282 academic publications and inductively develops a framework for digital transformation comprising eight foundational elements. The framework emphasizes digital transformation as a process wherein digital technologies induce disruptions, necessitating strategic responses from enterprises. These answers seek to optimize value development strategies while addressing structural modifications and organizational obstacles. The results of this process may be both advantageous and detrimental. Nadkarni and Prügl (2021) assert that traditional research often employs linear models to elucidate the relationships among numerous factors pertinent to digital transformation. They contend that digital transformation is inherently non-linear because of the intricate interconnections among various components, including technology, individuals, and organizations. Egodawele et al. (2022) examined 174 research articles published from 2013 to

2021 in thirty esteemed journals selected by the Australian Business Deans Council (ABDC). Their research indicated that digital transformation can be categorized into distinct components (see Verhoef et al., 2021). They found four significant theoretical constructs in the literature: resource-based view, dynamic capabilities, configuration theory, and ambidexterity theory. The authors endeavor to integrate various theories to establish a cohesive framework for comprehending the complex dynamics of digital transformation (Egodawele et al., 2022, pp. 5-8). Although they provided a thorough framework for digital transformation, they failed to explicitly define digital resources, digital capabilities, or the execution of a digital transformation plan. Kraus et al. (2021) identified three categories of digital transformation: technological, institutional, and sociological, in a thorough study. The emphasis was on the effects of digital transformation rather than on delineating digital transformation and its potential. Korhonen and Gill (2018) seek to assess digital capability and propose a framework delineating six levels and seven dimensions of digital capability. They examine the various tiers of digital proficiency, ranging from fundamental literacy to advanced leadership and transformation. There exist seven dimensions, spanning from zero to adaptable, each possessing distinct features. This study, however, exhibits a deficiency in conceptual clarity about the differentiation of various levels and stages. Nadkarni and Prügl (2021) examined several facets of digital transformation by reviewing 58 research published from 2001 to 2019, concluding that technology and stakeholders are the primary catalysts of transformation. They examine the fusion of digital technologies with organizational transformation efforts in the technology-centric realm. The actor-centric dimension highlights the significance of middle management in facilitating digital transformation, addressing any skills gaps, and evaluating the benefits and drawbacks of partnering with start-ups and purely digital enterprises to expedite transition. In the technology-

focused realm, they investigate the amalgamation of digital technologies and organizational transformation efforts. This study aims to elucidate digital transformation by emphasizing the concept of a digital ecosystem.

Theoretical Framework:

Warner and Wager (2019) define “digital capabilities” as an organization’s proficiency in utilizing digital technologies, including big data, cloud computing, the Internet, and artificial intelligence (AI), to foster innovation and facilitate commercial growth. We elucidate the notions of digital integration capability and digital platform capability, as they are significant catalysts of digital transformation.

Digital Integration Capabilities: The term “digital integration capabilities” denotes a company’s ability to effectively assimilate digital technology into its current systems and processes. This capability is crucial for the synchronization of digital tools, data, and workflows, which is necessary to attain enhanced efficiency and productivity (Xiong, et al, 2024).

Digital platform capabilities denote a company’s ability to create, manage, and utilize digital platforms. Platforms facilitate the establishment of ecosystems, promote collaboration, foster creativity, and generate value. Social media platforms, cloud-based services, and online marketplaces exemplify social media platforms (Bughin, Catlin, and Dietz, 2019; Popova et al., 2024).

Digital skills encompass the competencies, knowledge, and expertise required to utilize digital technologies for fostering innovation and facilitating organizational transformation. These competencies empower individuals and organizations to adapt to the rapidly evolving digital environment and encompass both technical and interpersonal capabilities (Marhraoui, 2023). It encompasses both technical expertise and business acumen, specifically the capacity to collect,

analyze, and assess data to make informed decisions, as well as an understanding of the strengths and limitations of Artificial Intelligence (AI) and Machine Learning in operational automation. Additionally, it requires knowledge of business strategy and its scope, along with the ability to collaborate effectively with cross-functional teams to further business objectives.

Equipped with the above three competencies, a corporation can leverage them to foster creativity, evaluate new concepts, and establish benchmarks for digital transformation. The digital capabilities are also referred to as digital innovation capabilities. To sustain a competitive advantage, it is essential to foster innovations by leveraging digital integration and platforms such as artificial intelligence, the Internet of Things, and data analytics. Consequently, the micro-foundation of digital transformation encompasses a comprehension of and a focus on cultivating the three essential competencies that are pivotal to digital transformation. To achieve digital transformation, firms must adeptly navigate the digital landscape and thrive in a perpetually evolving environment.

Digital Ecosystems:

A robust digital ecosystem is comprised of digital integration capabilities, digital platforms, and digital competencies. Digital integration competence refers to an organization's proficiency in seamlessly incorporating digital technologies into its existing systems and processes. This feature ensures the synchronization of various digital tools, data, and workflows to enhance productivity and efficiency. Digital platforms serve as the essential infrastructure that enables the trade of digital products and services (Adner, 2017). These platforms facilitate collaboration, innovation, and value generation by enabling the construction, management, and utilization of digital ecosystems (Márton, 2022). Examples include online markets, social networking platforms, and cloud-based services. The abilities required to proficiently employ digital platforms and

technology are referred to as digital talents. These abilities encompass soft skills such as cooperation, change management, and strategic thinking, alongside technical skills such data analysis, software development, and cybersecurity. Digital skills empower individuals and organizations to effectively leverage digital integration platforms and capabilities within a digital ecosystem (Zhang, Wang, and Peng, 2024). Organizations can foster innovation, improve consumer experiences, and maintain a competitive advantage in the rapidly evolving digital market by developing and fostering these abilities.

However, it is essential to recognize that the elements of digital capabilities, including integration, platform, and digital skills, do not interrelate in a linear manner to foster an ecosystem. A digital ecosystem is an intricate network of interconnected digital instruments. These digital tools encompass technical resources and digital platforms, together with engagement in enterprises and the services they provide. These digital tools collaborate to operate as a unified entity. A variety of technologies are employed in a robust digital ecosystem.

Digital platforms are essential, as they enable programmers to utilize application programming interfaces (APIs) to facilitate interaction and integration among diverse software.

A digital ecosystem facilitates the collaboration of businesses and partners to develop new goods or services that integrate effortlessly with existing offerings. Prominent platforms within this ecosystem encompass enterprise resource planning (ERP) software, customer relationship management (CRM) systems, and content management systems (CMS). Furthermore, digital ecosystems utilize instruments for project management, engagement, collaboration, and knowledge management. In recent years, India's digital financial ecosystem has experienced a significant transition, primarily attributable to the government's digital efforts. The nation has ascended as a global frontrunner in real-time payments, leveraging indigenous

technology innovations like the Unified Payments Interface (UPI). This innovation has placed India at the vanguard of the digital payments sector.

Digital Transformation:

A vibrant digital ecosystem is vital for facilitating digital change within and among enterprises. Digital transformation entails the application of digital technologies to innovate, enhance customer experience, and optimize processes. To create major transformations in operations and value generation, digital solutions must be incorporated into every aspect of an organization. This transformation needs the implementation of new technology, the reevaluation of business models, the development of an innovative culture, and the optimization of processes for enhanced agility and responsiveness (Adner, 2017). Embracing digital transformation enables firms to maintain competitiveness by enhancing their capacity to manage the intricacies of the digital landscape. A paper by Chung, Dietz, Rab, and Townsend from McKinsey and Company, published on September 11, 2020, indicates that digital ecosystems offer distinctive value by allowing participants to engage a wide array of clients from several industries. Moreover, members can augment or regulate engagement points along with the customer experience, relevant to both business-to-consumer and business-to-business contexts. Businesses do not depend exclusively on their own resources; rather, they form extensive collaborations with a diverse range of stakeholders, both within and outside conventional industry limits. These collaborative efforts foster dispersed innovation, augment productivity throughout value chains, better customer experiences, and generate new value opportunities for diverse stakeholders. Digital ecosystems, comprising networks of interlinked organizations, technologies, and services collaborating to provide holistic solutions, are crucial to this shift. Digital ecosystems leverage

synergies to provide firms with access to an expanded array of resources and capabilities.

Companies may provide cohesive, integrated solutions that generate value, foster innovation, and enhance customer experience through these ecosystems. Engaging in digital ecosystems enables organizations to expedite market entry, capitalize on partner capabilities, and attain operational advantages that would be challenging to realize independently. The symbiotic relationship between digital transformation and digital ecosystems significantly impacts the future of business and the continual adaptation of companies.

A hallmark of digital transformation is the network effects that facilitate positive feedback loops, propelling the evolutionary development of these ecosystems. Ecosystems provide goods and services that individual enterprises cannot create independently, hence attracting an increasing number of clients. The rise in clientele generates additional data, enabling artificial intelligence (AI) to enhance products, thereby optimizing operations and drawing in new customers. To serve customers. To establish a cohesive, customer-oriented value proposition that enables customers to acquire a diverse array of goods and services from a singular source, ecosystems integrate several stages of value chain creation. In addition to saving money, customers often acquire new experiences, which enhances their likelihood of remaining loyal.

Leading technology providers and vendors are promoting growth by aligning their strategies with ecosystem participants. They established data exchanges for ecosystem partners using application programming interfaces and supplied the necessary hardware and software to develop platforms. While “cloud solutions “as a service” have made it easier for partners to integrate partners into ecosystems, the expansion of 5G communications expansion is improving the speed and efficiency of interconnections. To improve the understanding and customer use of

ecosystems, advanced technology providers are delivering providing tools for managing large databases and utilizing leverage artificial intelligence. For example, there are several main factors that are going to undergo considerable changes in the ever-changing world of marketing: The pricing structure will evolve to become dynamic, adapting in real-time to fluctuations in market conditions, consumer behavior, and demand levels. This strategy enables firms to maximize their revenue and provide individual clients with pricing that is adapted to their specific needs. Access will be of the utmost importance, and businesses will be concentrating their efforts on making their goods and services accessible whenever and wherever customers require them from them. Because of this change, the convenience of the client will be improved, and traditional barriers will be removed. A significant number of products will be converted to digital formats, which reflects the increasing significance of technological advancement and innovation. This encompasses everything from pieces of software and digital material to goods and services that are only available online. Promotional techniques will be digital, utilizing advanced technology like artificial intelligence and augmented reality to create personalized and engaging marketing campaigns. Below, we provide a brief understanding of dynamic capabilities of firm and digital transformation/

Dynamic Capabilities of the Firm:

Teece, Pisano, and Shuen's (1997) research defines "dynamic capabilities" as a company's ability to "engage in adapting, integrating, and reconfiguring internal and external organizational skills, resources, and functional competences to match the requirements of a changing environment." Dynamic capabilities are employed to emphasize the integration of various unique capabilities to react swiftly and effectively to changes in the

external environment. The capability varies in correlation with the shifting focus of the business. Using the three managerial thinking of sensing, seizing, and transforming, capability can be made dynamic. As part of the sensing process, opportunities that exist outside the organization's boundaries are identified and evaluated. It is necessary to allocate resources effectively to take advantage of opportunities and maximize their potential. The destiny of organizations is being shaped. Below in Table 1 shows a framework of digital transformation and dynamic capabilities.

Table 1: Framework for Digital Transformation and Dynamic Capabilities

Dynamic Capability	Digital Transformation Aspect	Description
Sensing	Digital Sensing	Using digital tools and data analytics to identify opportunities and threats
Seizing	Digital Seizing	Leveraging digital technologies to capture opportunities and implement ideas into solutions.
Transforming	Digital Transformation	Reconfiguring resources and processes through digital means for staying competitive.

Again, we must understand that the relationship between dynamic capabilities and digital transformation is not always aligned in the same direction. Dynamic capabilities are related to an organization's ability to innovate and adapt while digital transformation is related to the integration of digital tools and technologies in business products and services that are entrenched in business strategy. An organization might have strong dynamic capabilities but not necessarily undergo a digital transformation, and vice versa.(Al-Moaid and Almarhdi, 2024). Based on the above description we can find 4 scenarios of organizations. High dynamic capability and high digital transformation, organizations which are innovative and digital leader. High dynamic

capability and low digital transformation, organizations which are traditional adapters. Low dynamic capability and high digital transformation, organizations which are technology reliant. Low dynamic capabilities and low digital transformation, organizations, conservative responder. Below we briefly define them.

Innovative and Digital Leaders: High dynamic capability and high digital transformation.

These organizations are at the forefront of innovation, effectively leveraging technology and adapting to market changes.

Traditional Adapters: High dynamic capability and low digital transformation. These organizations adapt well to market changes but use minimal digital technologies.

Technology Reliant: Low dynamic capability and high digital transformation. These organizations rely heavily on technology but may struggle to adapt to changes without strong dynamic capabilities.

Traditionalists: Low dynamic capability and low digital transformation. These organizations rely on traditional business practices and may struggle to compete in rapidly changing markets.

Conclusion

This article explains the principles of digital transformation through a paradigm centered on digital ecosystems. Digital ecosystems represent a multifaceted interconnection of digital platform capabilities, digital integration proficiency, and the digital competencies of people. Digital transformation enables enterprises to create, administer, and utilize digital platforms that promote collaboration, innovation, and value generation. Illustrative examples of these platforms encompass social media, cloud services, and online marketplaces, which constitute the basis for constructing interconnected ecosystems. These platforms facilitate the introduction of new

products and services, augmenting user engagement and fostering innovation in the digital era. By leveraging the capabilities of digital platforms, organizations can maintain competitiveness and adaptability in the dynamic market environment.

Digital competencies are crucial for business transformation to deal with evolving digital environment. These competencies include both technical expertise, such as data analysis, software development, and cybersecurity, as well as interpersonal skills like strategic thinking, collaboration, and change management. Armed with these competencies, individuals and organizations can adeptly utilize digital technologies to stimulate innovation and fulfill business goals. Digital competencies enhance the adoption and integration of digital platforms, allowing businesses to streamline processes, make informed decisions, and sustain a competitive advantage.

The incorporation of digital platform capabilities and digital skills within a digital ecosystem is essential for effective digital transformation. Digital ecosystems consist of interlinked enterprises, technologies, and services that cooperate to provide holistic solutions. By engaging in these ecosystems, organizations can tap into a wider array of resources and capabilities, expedite market entry, and attain operational efficiencies. The interplay among digital platforms, digital skills, and digital ecosystems establishes a solid groundwork for ongoing innovation, improves customer experiences, and maintains competitive advantage. We also provide an overview of dynamic capabilities and argue that digital transformation and dynamic capabilities are not always aligned with others, and they can provide different business strategies based on different sets of environments and digital needs.

References

- Adner, R. (2017). Ecosystem as Structure: An Actionable Construct for Strategy. *Journal of Management*, 43(1), 39-58.
- Al-Moaid, N.A.A., Almarhdi, S.G. (2024). Developing dynamic capabilities for successful digital transformation projects: the mediating role of change management. *J Innov Entrep* **13**, 85.
- Bughin, J., Catlin, T., and Dietz, M. (2019). The appropriate digital-platform strategy. *McKinsey Quarterly*, 2, 1-4.
- Burrus, D., and Mann, J. D. (2011). *Flash Foresight*. HarperCollins.
- Cennamo, C., Dagnino, G. B., Di Minin, A., and Lanzolla, G. (2020). Administration of Digital Transformation: Extent of Transformation and Methods of Value Co-Generation. *California Management Review*, 62(4), 5–16.
- Chung, V., Dietz, M., Rab, I. and Townsen, Z. (2020). Ecosystem 2.0: Climbing to the next level. *McKinsey Quarterly*, September 1-9.
- Egodawe, M., Sedera, D., and Bui, V. (2022). A Systematic Review of Digital Transformation. Literature from 2013 to 2021 and the formulation of a comprehensive apriori model for guidance. *Future research. arXiv preprint* arXiv:2212.03867.
- Gil-Gomez, H., Guerola-Navarro, V., Oltra-Badenes, R., and Lozano-Quilis, J. A. (2020). Customer Relationship Management: Digital Transformation and Sustainable Business Model innovation. *Economic Research-Ekonomska Istraživanja*, 33(1), 2733–2750.
- Hinterhuber, A., and Liozu, S. M. (2021). Digital transformation and the significance of customer-centricity Innovation: Interview with the Chief Value Officer of Thales in Managing

Digital. Transformation (pg. 186–191). Routledge.

Kohli, R. and Melville, N. P. (2019). Digital Innovation: A Review and Synthesis. *Information Systems Journal*, 29, Pages 200-223.

Kraus, S., Jones, P., Kailer, N., Weinmann, A., Chaparro-Banegas, N., and Roig-Tierno, N. (2021). Digital Transformation: A Comprehensive Examination of the Present Research Landscape. *Sage Open*, 11(3), 21582440211047576.

Korhonen, J. J., and Gill, A. Q. (2018, January). An Analysis of Digital Capability. In *ACIS 2018-29th Australasian Conference on Information Systems*.

Kreiss, D., and Brennen, J. S. (2016). Normative Models of Digital Journalism. *The Sage Handbook*. In the realm of digital journalism, pages 299-314.

Marhraoui, M. A. (2023). Digital Competencies for Project Managers: A Systematic Literature Review. *Procedia Computer Science*, 219, 1591-1598.

Márton, A. (2022). Progressing towards a digital ecology: ecological concepts for the examination of digital ecosystems. *Journal of Information Technology*, 37(3), 250-265.

Nadkarni, S., and Prügl, R. (2021). Digital Transformation: A Review, Synthesis, and Opportunities. For subsequent research. *Management Review Quarterly*, 71(2), 233–341.

Orji, C. I. (2019). Digital Business Transformation: Toward an Integrated Capability Framework for digitization and the development of commercial value. *Journal of Global commercial and Technology*, 15(1), 47–57.

Pereira, G. V., Luna-Reyes, L. F., and Gil-Garcia, J. R. (2020). Innovations in governance and digital Transformation and the creation of public value in Smart City programs. *Proceedings the 13th International Conference on the Theory and Practice of Electronic Governance*,

602 to 608.

Popova, P., Popov, V., Marinova, K., Petrova, M., and Shishmanov, K. (2024, June). The Digital Platform—novel opportunities and implementation strategy. In *2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 1-9. IEEE.

Pratt, M. K., and Boulton, C. (2023). Defining Digital Transformation: An Essential Disruption. Chief Information Officer. *CIO*, October 11.

Schwertner, K. (2017). Digital Transformation of Business. *Trakia Journal of Sciences*, 15(1). 388 -393.

Seymen, D. Ö. F. (2022). A customer-centric analytics framework and insights into digital. Transformation. *Knowledge Management and Digital Transformation Power*, pp 1-14.

Shrivastava, S. (2017). Digital disruption is transforming the customer experience: The digital Transformation strategy of telecommunications service providers. *Telecom Business Review*, 10(1), 41-52.

Teece, D.J., Pisano, G. and Shuen, A. (1997), Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509-533.

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., and Haenlein. M. (2021). Digital Transformation: A Multidisciplinary Reflection and Research Agenda. *Journal of Business Research*, 122, 889-901.

Vial, G. (2021). Comprehending digital transformation: A review and research plan.

Administration of digital transformation, *Managing Digital Transformation*, pages 13-66.

Zaoui, F., and Souissi, N. (2021, June). A framework for strategic digital transformation. In *2020 6th IEEE Congress on Information Science and Technology (CiSt)* 502-508.

Xiong, S., Paulraj, A., Dai, J., and Irawan, C. A. (2024). Asymmetry in digital integration capabilities and buyer product innovation: the dependent influences of environmental dynamism and inventive climate. *International Journal of Operations and Production Management*.

Zhang, X., Wang, P., and Peng, L. (2024). Formulating a Competency Model for Human Resource Directors (HRDs) in Exponential Organizations Experiencing Digital Transformation. *Sustainability*, 16(23), 10540.

Leveraging Machine Learning for Enhanced Supply Chain Decision-Making: A Study on Shipment Pricing and Mode Selection

Khushi Vijay Mehta (Kmehta6@pride.hofstra.edu)

Professor Alexander Pelaez

Hofstra University

1. Abstract

This study examines the dynamics of supply chain shipment price, with a particular emphasis on the interactions among freight rates, shipment methods, and national characteristics. By employing advanced statistical techniques such as Support Vector Machines (SVM), multinomial regression, and clustering, the study finds trends and factors that affect vendor selection. According to the multinomial regression model, freight prices are a significant factor in determining the mode of transportation. Truck and ocean shipments are found to be more cost-effective options, while air and air charter modes have a negative correlation with cost. While the SVM model offers predictive insights into vendor Incoterm selection, ensuring that decisions are in line with logistical and regulatory considerations, clustering analysis indicates geographical trends in shipment preferences influenced by infrastructure and economic variables. The results offer practical insights into cost-cutting techniques, vendor negotiation tactics, and compliance optimization while illuminating vendor behaviors that are rarely examined in the literature. By filling in the gaps in the literature, this study advances our knowledge of data-driven decision-making in global supply chains and provides useful suggestions for improving strategic planning and operational effectiveness.

2. Introduction-

Global supply chains are getting more complicated, presenting the need for advanced analytical methods to maximize price and logistics decision-making. International trade agreements, fluctuating transportation costs, and regulatory restrictions provide major obstacles for the pharmaceutical business when it comes to shipment pricing and mode selection. In supply chain logistics, traditional pricing models have mostly depended on market-based or cost-plus approaches, which frequently fall short in capturing the fluctuating nature of freight prices and vendor-specific circumstances. Comprehensive pricing models that successfully include shipping characteristics, vendor policies, and national economic considerations into predictive frameworks are still lacking, despite the growing popularity of data-driven decision-making. This study attempts to close this gap by improving shipment pricing and mode selection tactics

through the use of machine learning (ML) techniques. Big data has led to a growing use of machine learning algorithms for supply chain optimization. However, most of the current machine learning research in logistics has concentrated on inventory control and demand forecasting, leaving vendor incoterm selection and shipment price largely unexplored. This study closes a significant research gap in comprehensive pricing models by using Support Vector Machines (SVM) and multinomial regression to examine vendor incoterm selection and shipment pricing in pharmaceutical logistics.

3. Historical Background-

Over the past three decades, there has been a major evolution in the way data analytics are integrated into supply chain decision-making. Supply chain management has historically optimized price, logistics, and mode selection using deterministic models and rule-based decision algorithms. However, more sophisticated analytical techniques are now required due to the growing number of data, shifting economic situations, and the intricacy of international trade. To assess supply chain costs and improve logistics, scholars and professionals started using statistical models like decision trees and linear regression in the early 2000s. The shortcomings of these models, including their incapacity to represent high-dimensional data interactions and nonlinear correlations, were gradually revealed. As a result, machine learning (ML) approaches were adopted, which provide improved prediction capabilities and the capacity to reveal hidden patterns in intricate datasets.

The use of machine learning (ML) in supply chain logistics began to gain popularity in the 2010s, as research showed that algorithms like support vector machines (SVM), clustering techniques, and deep learning were beneficial for tasks including shipping pricing, demand forecasting, and route optimization. Regulatory restrictions, temperature-sensitive logistics, and the requirement for on-time delivery have presented special difficulties for the pharmaceutical supply chain, which has increased demand for AI-driven decision-making models. ML usage has been further boosted by the increasing availability of big data tools and computational resources, which allow businesses to combine historical shipment data with outside variables like trade regulations, economic indicators, and geopolitical threats. This change in history emphasizes how crucial it is to use machine learning (ML) to improve supply chain decision-making, especially when it comes to things like shipment pricing and mode selection.

4. Literature Review-

Shipment pricing is influenced by a wide range of factors, including vendor-specific terms like Incoterms, macroeconomic conditions, trade policies, and product features like weight, volume, and value. Various studies have explored these dimensions, employing diverse methodologies to extract meaningful insights into pricing and mode selection.

Particularly in pharmaceutical logistics, studies such as Zhu et al. (2021) demonstrate the importance of lead time, freight cost, and vendor procedures in determining shipping costs. Singh and Kapoor (2021) investigated the effects of lead time, freight cost, and shipment mode on incoterm selection in pharmaceutical supply chains using multinomial regression. Although machine learning techniques are not incorporated for improved predictive skills, this research emphasize the significance of standardized vendor agreements.

Wang et al. (2019) highlighted regional cost-saving opportunities by classifying nations based on shipment frequency and logistics costs using k-means clustering. In a similar vein, Kim et al. (2020) used hierarchical clustering to examine vendor behavior and found that sellers in similar regions varied in price.

The effectiveness of Support Vector Machines (SVM) in classifying modes of transportation according to product features was shown by Zhang et al. (2020), who achieved excellent predicted accuracy. Chen and Lee (2022) provided more evidence of SVM's resilience while dealing with unbalanced datasets, which is a frequent problem in shipment price studies. However, rather than offering practical supply chain methods, this research mostly concentrates on forecast accuracy.

According to Hosmer et al. (2013), multinomial regression has drawbacks, especially the presumption of linear connections between variables, which can oversimplify actual shipment decisions. Yet, multinomial regression is still a popular technique for examining categorical shipment modes in supply chain logistics (Singh & Kapoor, 2021).

Patel and Rao's (2022) research highlight the intricacy of pharmaceutical logistics, where shipment decision-making is heavily influenced by elements such as cold chain requirements, strict compliance standards, and emergency supply needs. According to these studies, incorporating operational and regulatory restrictions into shipment price models may improve decision-making efficacy and cost efficiency.

5. Gaps in the literature-

Even while we have gained a much better understanding of shipment pricing, there are still some substantial gaps. Prior studies often focus on single aspects of pricing, such as vendor processes or transportation options, without integrating these aspects into a comprehensive model. Furthermore, the dynamic interconnections between economic, geographic, and regulatory elements that affect shipping pricing and mode selection are not taken into consideration in most current studies. The challenges that the pharmaceutical business faces, such as regulatory restrictions, temperature-sensitive logistics, and the need for prompt choices to guarantee product integrity and timely distribution, are also not well studied. The inadequate use of advanced analytical methods to thoroughly examine shipment data represents another important gap. Although traditional methods offer useful insights, they frequently fail to recognize the multifaceted character of supply chain dynamics, which include interdependencies across elements like vendor preferences, freight costs, and infrastructure specific to a given nation. Furthermore, little study has looked at how shipment price methods fit in with suppliers' strategic goals, like cost reduction and legal compliance.

This paper fills these gaps by employing a comprehensive approach that combines multinomial regression, SVM, and clustering to examine shipment price in a worldwide pharmaceutical supply chain. This strategy not only looks at the discrete elements that affect prices, but it also uncovers intricate connections and trends that conventional approaches could overlook. This study intends to offer practical insights into enhancing shipment pricing and decision-making in the pharmaceutical supply chain by using these cutting-edge methodologies.

6. Methodology-

6.1. **Data Description:**

So based on the research gap identified at the end of the Literature review, this study presents the data driven model where it tries to predict various variables in a Global Omnichannel Pharmaceutical Supply chain. The case is represented by a dataset that contains logistics records of the shipment of pharmaceutical goods between the years 2006 to 2015. The data is obtained from a global repository of pharmaceutical supply chain records. There are 34 attributes, with 10,324 data points in the dataset. The values of features mix practically all data kinds, with integer, floating-point, and string values. The medications are sent to 43 customer nations across four continents from records. There are mainly-two supply chains: direct drop

shipping and regional distribution centers (RDCs)-based shipment, which relies on a distribution network ([Zhu et al., 2021](#)) to provide pharmaceutical items.
Information about variables-

6.2. Variable Categorization:

We split the 34 attributes into two categories: qualitative and quantitative. Country, Fulfil Via, Vendor Incoterm, Shipment Mode, Product Group, Sub Classification, Vendor, Manufacturing Site, First Line Designation, medical information linked factors, and Identification associated variables are qualitative variables. Lead Time, Line-Item Quantity, Item Value, Pack Price, Unit of Measure, Pack Price, Unit Price, Weight, Freight Cost, and Line-Item Insurance are the quantitative variables.

6.3. Clustering-

Based on important characteristics like vendor, Incoterm, product group, dosage form, and unit price, the clustering approach is utilized to find pricing trends among vendors. The following five clusters were found: “Economy,” “Premium,” “Standard,” “Specialty,” and “Diverse.” These clusters show clear patterns in the features and costs of the products:

Economy: Vendors that offer simpler product features and lower unit prices to appeal to budget-conscious consumers.

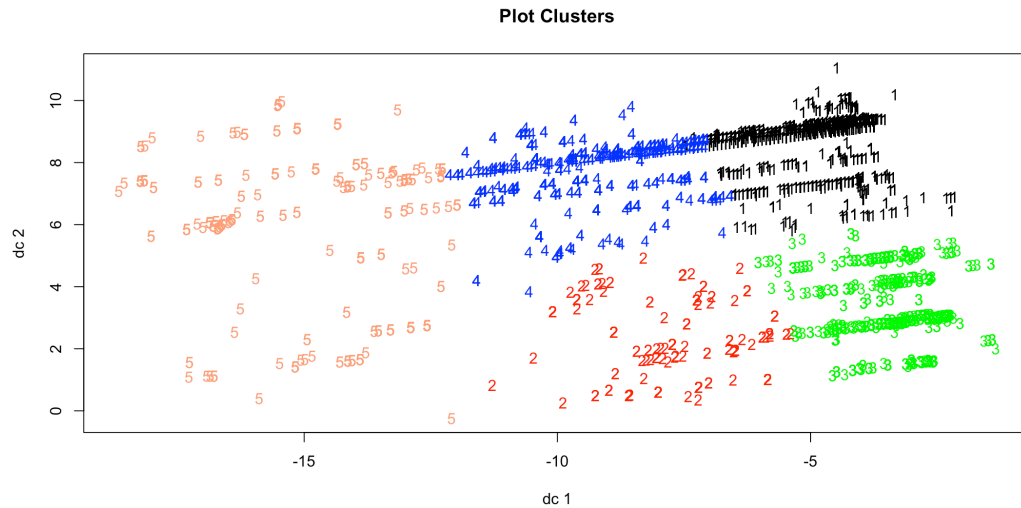
Premium: Usually connected to high-end goods or specialized services, these vendors charge noticeably more per unit.

Standard: A fair pricing and product selection strategy that meets the expectations of typical consumers.

Specialty: Suppliers that focus on specialty or niche pharmaceuticals, with more expensive units and unique dose forms.

Diverse: Vendors that exhibit flexibility in their market positioning through a variety of price tactics.

The most obvious variances are seen in clusters like “Economy” and “Premium,” which emphasize how vendor-specific pricing tactics affect transportation expenses. These observations highlight how crucial vendor tactics and Incoterm selection are to cost control and supply chain efficiency.



Name of the clusters

“Economy”, “Premium”, “Standard”, “Specialty”, “Diverse”

Result-

Group.1 vendor.inco.term product.group vendor dosage.form unit.price

Group		Vendor.inco.term	Product.group	vendor	Dosage form	Unit Price
1	1	7.331052	3.015803	58.78087	10.122678	15.00472
2	2	5.630000	3.143333	12.71667	9.810000	67.34667
3	3	5.435739	2.974570	17.01718	9.525086	14.05430
4	4	6.039949	3.515536	54.73050	13.579581	74.03678
5	5	5.837237	3.906323	56.07963	15.618267	140.48595

6.4. Multinomial Regression-

This method looks for patterns in the frequency and mode of shipments by country, as well as the correlation between these patterns and freight rates. The study employed a multinomial logistic regression model to evaluate the relationship among country, freight charges, and shipping mode. The primary objectives were to identify areas for cost reduction and obtain practical knowledge about supply chain operations.

The following significant coefficients from the multinomial regression analysis show how important freight prices are when deciding on a mode of transportation:

- **Air Mode:** In line with the idea that air shipments are more costly, the likelihood of using this mode drops by 0.42 units for every unit increase in freight charges. Air Mode (Negative Preference): Zimbabwe (-0.74) and Ethiopia (-0.67) show a negative preference, most likely because of financial constraints or poor airport infrastructure.
- **Air Charter:** Given its premium status, the negative coefficient of -0.78 indicates a larger inverse relationship with higher freight prices. Air shipments are commonly used by countries with higher economic activity or urgent delivery needs, such as Guatemala (1.05) and Lebanon (0.95).
- **Truck Shipment and Ocean Modes:** These modes have positive coefficients of 0.56 and 0.74, respectively, which show that they are associated with lower freight charges and are economical choices for large shipments. Truck Shipments are predominate in countries like Chad (1.0), South Sudan (1.2), Rwanda (0.9), Nigeria(0.85) and Uganda (0.88), which have better-connected road networks or are inland. Ocean Shipments are ideal for long-distance deliveries and bulk goods to countries with port access, like Thailand (1.02) and Kenya (0.9).

The country-level impacts are significant and show how economic and geographic factors affect the preferences for shipment modes

Regression Coefficients:

Variable	Air	Air Charter	Ocean	Truck
Intercept	12.91	-2.21	1.12	1.35
Freight Cost (USD)	-0.00012	-0.0000968	-0.000437	0.000121
Country Angola	0.6386	-0.9359	-2.8702	10.4439
Country Ghana	2.5208	1.1464	12.1189	11.9503
Country Nigeria	-9.6866	5.4251	-3.2367	-4.78299

This study sheds light on vendor preferences and behaviors, which are rarely examined in the literature, and how they affect the selection of Incoterms. For instance, vendors often select risk-averse Incoterms like Delivered Duty Paid when delivering to markets with complex regulations, which reflects their operational objectives.

By showing how Incoterm selection varies according to important factors, the model provides practical insights into:

Cost-cutting measures by aligning shipment mode with freight cost efficiency. Vendor negotiation tactics that consider country-level shipment preferences and infrastructure constraints. Compliance optimization by selecting Incoterms suited for regulatory environments.

“The model’s overall performance and goodness of fit are indicated by the residual deviance of 13,466.59 and the AIC value of 13,818.59. With lower values signifying a better trade-off between complexity and fit, the AIC aids in comparing various models. As a measure of the model’s unexplained variance, the residual deviance indicates how well the model explains the observed data; lower values indicate a better match.

These findings address a gap identified in the previous literature and pave the way for future research on vendor decision-making in supply chains.

6.5. **Support Vector Machine-**

I have examined the elements that affect vendor selection incoterm using the Support Vector Machine approach.

Support Vector Machine is suitable for this analysis because the dependent variable vendor incoterm has eight various classes, representing different vendor contracts. SVM with a linear Kernel method effectively handles multiclass classification problems, especially when the data points are linearly separable or nearly, so in a transformed feature space. The model shows a high degree of correlation between the predicted and actual classifications, with an overall accuracy of 98.61% and a kappa value of 0.9781. The model’s resilience, especially for less common incoterm classes, is further demonstrated by the balanced accuracy across classes.

Support Vector Machine provides useful information that can guide vendor agreements, policy decisions and shipment mode optimization. The dataset’s intrinsic imbalance is addressed by the addition of class weights, which guarantee that underrepresented incoterms are sufficiently considered throughout the classification process. For real-world applications, where choices on less common incoterms can have a big influence on supply chain efficacy and cost-effectiveness, this balance is essential.

Call:

svm(formula = vendor.inco.term ~ shipment.mode + vendor +

```
country, data = Path_SCM_df, kernel = "linear", cost = 1,  
class.weights = list(`1` = 10, `2` = 1, `3` = 10, `4` = 1,  
`5` = 1, `6` = 1, `7` = 1, `8` = 1), scale = TRUE)
```

Parameters:

SVM-Type: C-classification

SVM-Kernel: linear

cost: 1

Number of Support Vectors: 645

Confusion Matrix and Statistics

	Reference								
Prediction	1	2	3	4	5	6	7	8	
1	2	2	0	0	0	1	0	0	
2	0	22	2	0	0	3	26	5	0
3	0	6	9	19	5	5	0	0	
4	0	0	0	14	20	0	2	0	0
5	0	0	0	0	7	0	0	0	
6	1	38	0	4	0	27	38	13	0
7	0	7	0	0	0	6	37	9	0
8	0	0	0	0	0	0	5	40	4

Overall Statistics

Accuracy : 0.9861

95% CI : (0.9837, 0.9883)

No Information Rate : 0.5234

P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.9781

Mcnemar's Test P-Value : NA

Performance Metrics by Class

Metric	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
Sensitivity	0.6667	0.8073	1.0000	0.9841	0.4667	0.9856	0.9547	1.0000
Specificity	0.9997	0.9966	0.9966	0.9998	1.0000	0.9926	0.9987	1.0000
Pos Pred Value	0.4000	0.8672	0.2045	0.9986	1.0000	0.9800	0.9668	1.0000
Neg Pred Value	0.9999	0.9947	1.0000	0.9974	0.9992	0.9947	0.9982	1.0000
Prevalence	0.0003	0.0266	0.0009	0.1398	0.0015	0.2691	0.0385	0.5234
Detection Rate	0.0002	0.0215	0.0009	0.1375	0.0007	0.2652	0.0367	0.5234
Detection Prev.	0.0005	0.0248	0.0043	0.1377	0.0007	0.2706	0.0380	0.5234
Balanced Accuracy	0.8332	0.9019	0.9983	0.9919	0.7333	0.9891	0.9767	1.0000

The following are the findings from the Support Vector Machine:

Key Features: According to the investigation, the country, vendor, and shipment mode all have a big impact on the choice of incoterm. These characteristics capture the logistical limitations, operational preferences, and regulatory considerations that influence choices. Although the accuracy is generally great, some classes (such as class 1 and class 5 have lesser sensitivity, indicating areas where prediction accuracy could be increased through feature engineering or more data representation refinement. Supply Chain Managers can adjust their operations to match ideal incoterms by utilizing SVM's predictions, that lowers expenses, boosts productivity, and guarantees adherence to local laws. Significant gains could be made, for example, by standardizing procedures across vendors and nations or concentrating on increasing the effectiveness of distribution modes linked to underperforming incoterms.

Future Directions-

Despite the SVM model's remarkable performance, there is room for improvement. The model's capacity to manage intricate interactions may be improved by hyperparameter tuning, adding more predictive features (such as freight costs, product weight, and delivery urgency), and investigating different kernels.

7. Pre-COVID vs. Post-COVID: Pharmaceutical Supply Chain Analysis

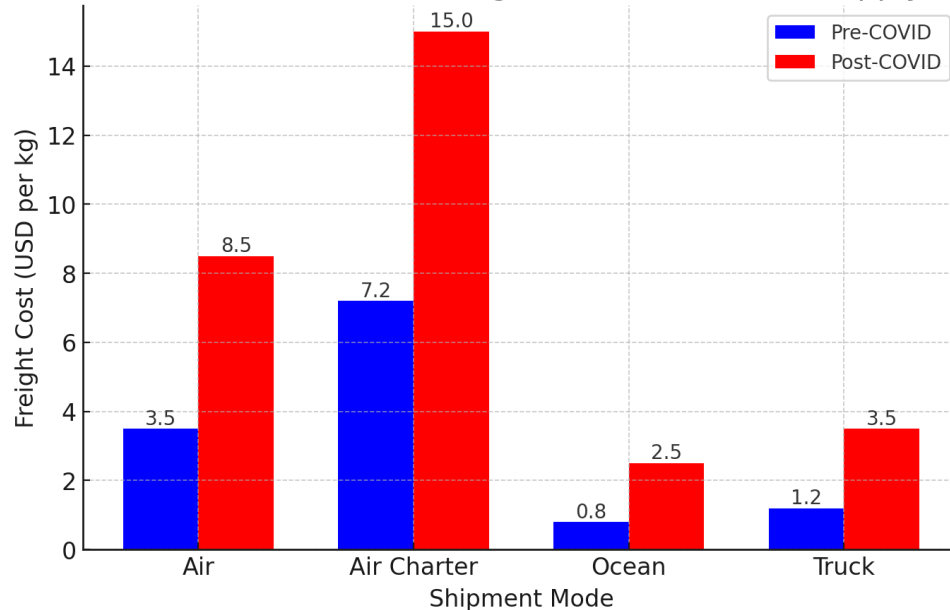
7.1. Pre-COVID-

Just-in-Time (JIT) inventory solutions were used to reduce costs in the pharmaceutical industry's rather reliable supply chains. China and India were the primary suppliers of generic medications and Active Pharmaceutical Ingredients (APIs). The cost of shipping remained constant, and high-value and urgent medications were mostly shipped by air freight. Stockpiling was restricted to necessary medications, and demand followed predictable seasonal and regional trends. Cross-border pharmaceutical exports were less restricted, and trade rules were more liberal.

7.2. Post-COVID-

The pandemic caused transportation bottlenecks, shortages of raw materials, and factory closures. Reduced airline capacity caused air freight charges to soar, while port congestion and container shortages affected ocean freight. To lessen dependency on a small number of supply countries, many nations began encouraging local pharmaceutical manufacture. There were shortages of essential medications and APIs because of export restrictions enforced by nations. Increased demand for necessary medications (such as vaccines and antivirals) led to supply imbalances and stockpiling. AI-driven supply chain models have been used by pharmaceutical corporations to forecast demand and streamline operations.

7.3. Graph: Pre-COVID vs. Post-COVID Freight Costs for Pharmaceuticals-
Pre-COVID vs. Post-COVID Freight Costs in Pharma Supply Chain



8. Conclusion-

This study aimed to identify the variables affecting shipment pricing, assess trends, and suggest workable optimization solutions by thoroughly analyzing a solid dataset that spans a decade of

data. The results highlight how crucial it is to use cutting-edge analytical methods to tackle the complexity of contemporary supply networks. The pharmaceutical industry's shipment pricing research highlights the complexity of cost determinants, which include method of transportation, geography, product characteristics, and vendor-specific policies such as incoterms.

Moreover, macroeconomic elements like currency volatility and trade restrictions make price decisions even more difficult. Using the three approaches, this study was able to address a variety of aspects of the issue, from policy-oriented analysis and predictive modeling to exploratory data segmentation. Logistics planning can be streamlined with the capacity to forecast incoterm and transportation method preferences, guaranteeing cost effectiveness without sacrificing service quality. By identifying actionable trends in shipment frequency and pricing by location, clustering allowed for the development of customized strategies for high-cost or underperforming clusters. Policymakers can more efficiently allocate resources by using multinomial regression insights, which shed light on the relative significance of factors like freight cost and shipment mode.

Lastly, in an increasingly interconnected world, companies can ensure the smooth delivery of essential items by implementing a data-driven approach to create supply chain systems that are more robust and efficient. Future studies should also investigate ways to optimize temperature-sensitive logistics, guarantee adherence to changing rules, and reduce the risks associated with time-sensitive deliveries, considering the particular difficulties the pharmaceutical business faces. Aligning shipment price methods with stakeholders' strategic goals—such as cost-effectiveness, sustainability, and product integrity—should also be a top priority. These models could be further refined to provide actionable insights catered to various markets and shipment modes through cooperative research with industry stakeholders that makes use of big data and real-time analytics.

9. References

Dataset reference: https://data.usaid.gov/HIV-AIDS/Supply-Chain-Shipment-Pricing-Dataset/a3rc-nmf6/about_data

- Chopra, S., & Meindl, P. (2022). *Supply Chain Management: Strategy, Planning, and Operation*. Pearson.
- Fisher, M. L. (1997). What is the right supply chain for your product? *Harvard Business Review*, 75(2), 105-116.\

- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer.
- Zhang, X., & Shen, W. (2021). Machine learning applications in supply chain management. *International Journal of Production Research*, 59(16), 4895-4915.
- Yao, Y., & Liu, W. (2020). Freight cost modeling in global supply chains. *Journal of Business Logistics*, 41(3), 205-224.
- **Christopher, M. (2016).** *Logistics & Supply Chain Management (5th ed.)*. Pearson.
- Zhu, S., Song, L., Hazen, B. T., Lee, K. L., & Cegielski, C. G. (2021).

HUMANITARIAN LOGISTICS NETWORK SYSTEM DESIGN USING MACHINE LEARNING BASED ON THE PRODUCTIVITY-DRIVEN CLUSTERING EVALUATION METHOD

Jae-Dong Hong, Industrial Engineering, South Carolina State University,
Orangeburg, SC 29117, 803-536-8861, jhong@scsu.edu

ABSTRACT

The humanitarian logistics network system (HLNS) design has become vital due to recent global natural or weather-related extreme events/disasters, including pandemic outbreaks. The *HLNS* design problem has recently required multi-objective models and has been evaluated using the data envelopment analysis (DEA)-based methods. This paper proposes a machine learning algorithm (MLA) based on the productivity-driven clustering (PDC) method for assessing various *HLNS* configurations. The results between the conventional DEA methods and the proposed MLA are compared through a case study. We demonstrate that the proposed design framework would help the decision-makers better evaluate the productivity of *HLNS* configurations and identify the productive and robust ones.

Keywords: Humanitarian logistics network system, Data envelopment analysis, Machine learning algorithm, Productivity-driven clustering method

INTRODUCTION

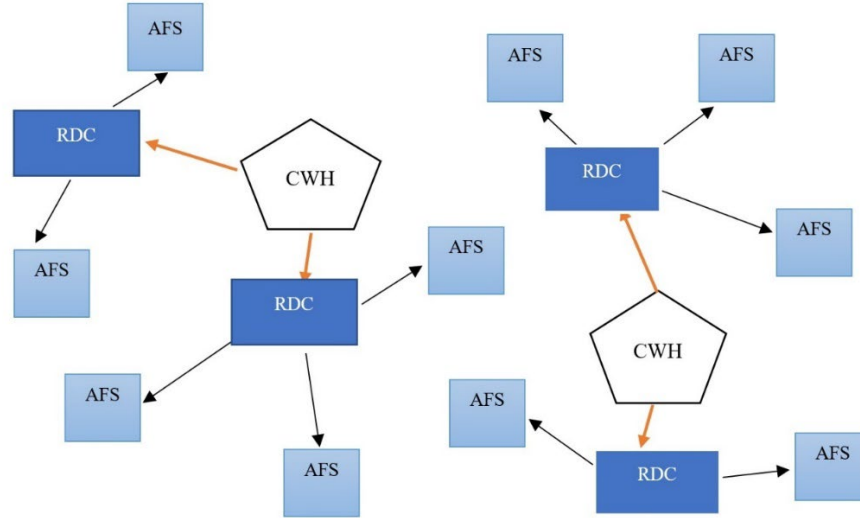
The recent weather-related extreme events or disasters and pandemic outbreaks have made the design of the humanitarian logistics network system (*HLNS*) a critical strategic issue for the world in a pre-disaster scenario. The *HLNS* design problem is locating and allocating disaster response facilities (DRFs). Creating and maintaining the productive and effective *HLNS* is essential in providing vaccines or relief items such as food, first aid, drinking water, and daily commodities to alleviate people's suffering (Day et al., 2012; Boonmee et al., 2017; Shavaranim, 2019; Hong, 2023). We aim to propose and demonstrate how to design a productive *HLNS* configuration under the risk of DRF disruptions. For this purpose, consistent and robust DRF performance evaluation is essential.

The *HLNS* considered in this paper is a two-echelon supply chain system with three distinctive DRFs, as shown in Figure 1. They are (i) Central Warehouses (*CWHs*), where humanitarian items are stored; (ii) intermediate response facilities termed Relief Distribution Centers (*RDCs*), where disaster-affected people can gain access to relief goods productively; (iii) Affected Sites (*AFSs*) that would need humanitarian items. The key objective of the strategic level is to strengthen emergency preparedness. Selecting the most cost or distance-effective location of *RDCs* and *CWHs* among a set of candidate locations, establishing the distribution of humanitarian supplies throughout the *HLNS*, and assigning *AFSs* to *RDCs* and *RDCs* to *CWHs* would be essential. However, traditional cost-based DRF location-allocation models implicitly assume that all DRFs will always provide service or are available and do not consider an associated risk of facility disruption. Due to natural and extreme weather-related disasters, pandemics, and/or accidents, all DRFs are open to disruptions. Such disruptions would be worsened due to a lack of flexibility and interdependency, commonly presented in the general supply chain systems.

Evaluating various *HLNS* alternatives and identifying the most productive schemes would be critical for logistics network planning. The multi-objective programming (MOP) models request decision-makers to assign different weights to the deviational variable of each objective. Decision-makers must reflect on the importance and desirability of deviations from the multiple objectives. However, the productivity or efficiency of the *HLNS* generated by solving the model with a fixed weight is unknown. No typical procedure for assigning values to the weight factors is available to ensure that decision-makers will find the most desirable solution to an MOP problem. Ragsdale (2022) suggests that an iterative procedure

should be followed, using a particular set of weights, concluding that we must repeat this procedure multiple times to find the most needed solution for decision-makers. Consequently, it is inevitable for decision-makers to use some of their subjective judgment. Evaluating various *HLNS* network schemes objectively, not subjectively, and selecting the most productive alternatives would be essential for designing *HLNS*. Hence, the researchers should be able to propose a better procedure for assessing the efficiency of all options generated by the MOP model and select the most desirable one(s) without any subjective judgment.

Figure 1: Two-echelon humanitarian logistics network system



Data envelopment analysis (DEA) was developed to get decision-making units' relative efficiency or productivity (DMUs), which produce multiple outputs using multiple inputs. The traditional DEA (T-DEA), proposed by Charnes et al. (1978), yields a single, comprehensive performance measure for each DMU. With the principle of maximizing efficiency, the T-DEA allows each DMU to be evaluated based on its most favorable weights. It is called self-evaluation. Consequently, under the T-DEA model, each DMU is allowed to disregard unfavorable outputs/inputs to maximize its self-efficiency. Thus, the lack of discrimination turns out to be the most critical weakness of T-DEA.

Out of many DEA-based methods to remedy this deficiency of T-DEA, the cross-efficiency (CE) DEA method is the most widely applied (Liu et al., 2019). The main idea of peer evaluation rather than T-DEA's self-evaluation makes CE-DEA an essential tool for assessing DMUs. As Doyle and Green (1994) note, the primal issue of applying the CE-DEA model is the non-uniqueness of cross-efficiency scores (CESSs) due to the often-present multiple optimal DEA weights. Another critical issue is that the CE method, significantly affected by the DMUs to be assessed, tends to generate inconsistent rankings.

Machine learning (ML) means instructing a machine or an algorithm to perform tasks. It is a subfield of artificial intelligence, an automated process extracting patterns from data. Malone et al. (2020) state, "The function of an ML learning system can be descriptive, predictive, and or prescriptive." In other words, the ML learning system uses the data to describe what happened, the data to expect what will happen, or the data to make recommendations about what action to take. Kelleher et al. (2015) state, "Machine learning algorithm (MLA) automates the process of learning a model that captures the relationship between the descriptive features and the target feature in a dataset." MLA works by searching through a set of possible

prediction models for the model that best captures the relationship between the descriptive features and target features in a dataset. When using ML, analysts begin with a model to work with, supply the data, and let the computer model train itself to find patterns and/or make predictions. Over time, the analyst may tweak the model to improve its performance (Kultur & Calayan, 2017). Zhu et al. (2021) propose combining DEA with MLA to predict the DEA efficiency of new DMUs by discussing four ML-DEA algorithms. Wu and Wu (2022) classify the DEA scenarios into overfitting and underfitting scenarios and propose a mean score methodology based on MLAs. They insist that the proposed method advances the development of DEA in the context of big data and its integration with ML. Both methods that Zhu et al. (2021) and Wu and Wu (22) propose are based on DEA, which has shown severe weaknesses, let alone excessive computational times.

Cluster analysis or clustering is an exploratory data analysis tool for solving classification problems. Clustering aims to identify meaningful segmentations or groupings of entities within a data set, focusing on distance-based analysis. In other words, clustering sorts entities into clusters or groups so that the degree of association is high among members of the same cluster and low among members of different clusters (see Delen, 2021). One of many tools for calculating distance measures is the centroid-based technique. This paper modifies the original cluster analysis following the productivity formula and proposes an ML based on the productivity-driven clustering (ML-PDC) method. What distinguishes the proposed approach from other clustering analysis methods is that the ML-PDC utilizes the overall minimum and maximum points to represent the clusters rather than the centroid of each cluster. The ML-PDC method divides all DMUs into productive and unproductive clusters and allows some DMUs to belong to both clusters. If a DMU's inputs are close to the minimum point, it would belong to the productive group regarding its inputs. Regarding outputs, the DMU belongs to the productive group if it is close to the maximum point. Without ignoring or sacrificing any given inputs/outputs, the ML-PDC method would eliminate the critical issues of DEA-based evaluation models and rate the DMUs more impartially and consistently than the DEA-based methods.

The significant contribution of this study is to propose a different framework consisting of formulating the design problem as an MOP model and evaluating and identifying the robust *HLNS* schemes for efficient and effective disaster/emergency relief operations. The ML-PDC method can allow us to evaluate various *HLNS* configurations objectively, not subjectively, and select the most productive alternatives, which would be essential for planning the *HLNS* schemes. This paper will attempt to apply the proposed method for designing *HLNS* for the first time and compare the results with those generated by CE-DEA, the most popular DEA-based method.

The research question is how to evaluate and consistently rank *HLNS* schemes generated by the mathematical programming models. What distinguishes the present paper is that the proposed ML procedure based on the PDC methods can identify the top-rated *HLNS* configurations, which could be different from the top-productive DMU that CE-DEA identifies. Thus, decision-makers can assess these high-ranked logistics networks together. The proposed approach can attract federal and local emergency/disaster response officials to develop a more flexible, robust, and productive emergency logistics network system.

FORMULATION OF THE *HLNS* MODEL

We follow the *HLNS* model that Hong and Jeong (2019) consider.

Sets:

- W : index set of candidate locations for *CWHs* ($w = 1, 2, \omega$)
- N : index set of AFSs n ($n = 1, 2, \dots, \mu$)
- J : index set of candidate locations for *RDCs* ($j = 1, 2, \mu, \mu+1, \dots, \mu+\omega$)

M : $M = \{N, I\}$, index set of AFS s and CWH s ($m = 1, 2, \dots, \mu, \mu+1, \dots, \mu+\omega$)

Parameters:

f_w : fixed cost for constructing and operating CWH_w
 θ_j : fixed cost for constructing and operating RDC_j
 a_{wj}^1 : shipping/transportation cost per mile per one unit of demand from CWH_w and RDC_j
 a_{wm}^2 : shipping/transportation cost per mile per one unit of demand from CWH_w and AFS_m
 d_{wj} : distance between CWH_w and RDC_j
 d_{jm} : distance between RDC_j and AFS_m
 C^{max} : maximum number of RDC s can be built
 CAP_w^{max} : capacity of CWH_w
 CAP_j^{max} : capacity of RDC_j
 h_m : demand of NBS_m (can be either AFS or RDC or CWH)
 W^{max} : maximum number of CWH s can be built
 q_w : minimum number of RDC s that CWH_w can handle
 Q_w : maximum number of RDC s that CWH_w can handle
 l_j : minimum number of AFS s that RDC_j can cover
 L_j : maximum number of AFS s that RDC_j can cover

Decision variables:

C_j : binary variable showing whether AFS_j is selected as RDC_j
 H_w : binary variable deciding whether a candidate CWH_w is selected
 ρ_{wj} : binary variable showing whether RDC_j is covered by CWH_w
 γ_{jm} : binary variable showing whether location m is covered by RDC_j
 z_{wjm} : binary variable showing whether location m is covered by CWH_w through RDC_j

Assumption (see Figure 1):

- (i) All AFS s and potential CWH locations are the candidates for RDC s to be located.
- (ii) Due to some realistic requirements, a CWH must only be located at one of the candidate CWH locations.
- (iii) An RDC cannot be located for any potential CWH location. An RDC should cover unselected CWH locations.
- (iv) A CWH feeds its own demand, and an RDC covers its own demand and demands from its covered AFS s.
- (v) When a DRF , CWH/RDC , is disrupted, it cannot satisfy any demand to be expected to cover.

The total logistics cost (TLC) for this problem set is given by

$$TLC = \left[\sum_{w \in W} f_w H_w + \sum_{w \in W} \sum_{j \in M} \left(\sum_{m \in M} h_m \right) a_{wj}^1 d_{wj} z_{wjm} \right] + \left[\sum_{j \in M} \theta_j C_j + \sum_{j \in M} \sum_{m \in M} h_m a_{jm}^2 d_{jm} \gamma_{jm} \right], \quad (1)$$

where

$$z_{wjm} = \rho_{wj} \cdot \gamma_{jm}. \quad (2)$$

The first goal is to minimize the TLC in (1), which is the classical objective of most facility location-allocation models. The second goal is to minimize the longest delivery distance between *CWHs*, *RDCs*, and *AFSs*, which is equivalent to minimizing the maximum coverage distance (*MCD*) such that an *RDC* covers each *AFS*, and a *CWH* covers each *RDC* within the endogenously determined distance. A long *MCD* will cause ineffectiveness in the resulting *HLNS*. Now, *MCD* is given by

$$MCD = \text{Max}\{d_{jm}\gamma_{jm}, d_{wj}\rho_{wj}\}, \forall w, j, \text{ and } m. \quad (3)$$

DRFs should be located at the least likely locations to be disrupted to enhance emergency supply chain resilience. The third goal is maximizing the expected amount of demands covered (*EDC*) by the *DRFs*, which is expressed through several algebraic manipulations as

$$EDC = \sum_{w \in W} \sum_{j \in M} \left[\sum_{m \in M} (1 - q_w)(1 - p_j)(z_{wjm}h_m) \right] + \sum_{w \in W} (1 - q_w)(h_w H_w), \quad (4)$$

where

q_w – the risk probability that the *CWH_w* is disrupted (or disruption probability).

p_j – the risk probability that the *RDC_j* is disrupted (or disruption probability);

In the *HLNS* design, it is reasonable that each location should be within a certain distance from the nearest *DRFs* to be served in a disaster/emergency. It would be conceivable to claim that the maximum effective coverage distance (*MECD*), denoted by φ_{cv} , should be shorter than *MCD* in (3). However, it would be desirable to maximize the covered demands within φ_{cv} , while minimizing *MCD*. The next goal is maximizing the covered demands in case of disaster, *CDE*, which is given by

$$CDE = \sum_{m \in M} \sum_{j \in J} h_m \kappa_{jm} \gamma_{jm} + \sum_{w \in W} h_w H_w \quad (5)$$

where a binomial parameter, κ_{jm} , is

$$\kappa_{jm} = \begin{cases} 1, & \text{if } d_{jm} \leq \varphi_{cv} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Let the nonnegative deviation variables, δ_{TLC}^+ , δ_{TLC}^- , δ_{MCD}^+ , δ_{MCD}^- , δ_{EDC}^- , δ_{EDC}^+ , δ_{CDE}^- , and δ_{CDE}^+ represent the amounts by which each goal value deviates from the target value. Then, let $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, $\sum_{\kappa=1}^4 \alpha_{\kappa} = 1$, denote relative weights assigned to the corresponding target or goal. We formulate the following weighted goal programming (WGP) model, where the objective function minimizes the total weighted percentage deviation:

$$\text{Min } Z_{\alpha} = \alpha_1 \frac{\delta_{TLC}^+}{TLC_{min}} + \alpha_2 \frac{\delta_{MCD}^+}{MCD_{min}} + \alpha_3 \frac{\delta_{EDC}^-}{EDC_{max}} + \alpha_4 \frac{\delta_{CDE}^-}{CDE_{max}}, \quad (7)$$

subject to

See Hong and Jeong (2019) for complete constraints.

Theoretically, the weight assigned to each objective should be any value between 0 and 1, but this study limits each weight to discrete values, changes between 0 and 1 with an increment of 0.1. Solving the above WGP model for a given set of weights generates one *HLNS* scheme with a group of optimal four-performance measures. Depending on the weight values, there will be various *HLNS* configurations. To evaluate these generated *HLNS* schemes, we apply CE-DEA and the proposed PDC method and compare the results of these two methods.

DATA ENVELOPMENT ANALYSIS METHODS

Traditional DEA (T-DEA)

Note that each DMU represents an *HLNS* scheme yielded by solving the WGP model given in (7)-(23) for a fixed value of each weight. Let E_t represent the DEA score for DMU_t , and DMU_t is said to be productive/efficient only if $E_t^* = 1$. Now, we formulate the following linear programming (LP) model of T-DEA with two outputs, *EDC* and *CDE*, and two inputs, *TLC* and *MCD*, as follows:

$$\text{Max } E_t = u_{1t}EDC_t + u_{2t}CDE_t, \quad (8)$$

subject to

$$v_{1t}TLC_t + v_{2t}MCD_t = 1, \quad (9)$$

$$(u_{1t}EDC_d + u_{2t}CDE_d) - (v_{1t}TLC_d + v_{2t}MCD_d) \leq 0, d = 1, \dots, D, \quad (10)$$

$$u_{1t}, u_{2t}, v_{1t}, v_{2t} \geq 0,$$

where

D = number of DMUs under evaluation

u_{rt} = multiplier/weight assigned by DMU_t to output r , $r = 1, 2, \dots, R$, and $t = 1, 2, \dots, D$

v_{it} = multiplier/weight assigned by DMU_t to input i , $i = 1, 2, \dots, S$

Cross efficiency DEA

The cross-efficiency (CE) method consists of self-evaluation (Phase 1) and peer evaluation (Phase 2). See Sexton et al. (1986). The first phase computes efficiency scores (Ess) using the model by (8)-(10). In the second peer-evaluation phase, the multipliers/weights generated in the self-evaluation phase are applied to all DMUs to get each DMU's cross-efficiency score (CES) as follows:

$$E_{td} = \frac{u_{1t}^*EDC_d + u_{2t}^*CDE_d}{v_{1t}^*TLC_d + v_{2t}^*MCD_t}, \quad t \text{ and } d = 1, \dots, D, \quad (11)$$

$$\bar{E}_t = \frac{1}{D} \sum_{d=1}^D E_{dt}. \quad (12)$$

PRODUCTIVITY-DRIVEN CLUSTERING (PDC) METHOD

The proposed approach, which evaluates each DMU's performance without applying DEA methods, differs from some DEA-based clustering methods like Po et al. (2009) and Chen et al. (2022). As T-

DEA classifies all DMUs under evaluation into two groups, separating productive DMUs from unproductive DMUs, there are only two (2) clusters in the proposed method. The first cluster is called a productive one, while the second is an unproductive cluster. Given a set of n DMUs with m inputs and s outputs, all DMUs under evaluation will be classified into one of these two clusters or both for input and output, respectively. What separates the proposed PDC method from other DEA-based clustering methods is that the global minimum and the maximum points are set to represent the clusters rather than the centroid-based technique.

Let c_{il}^k and c_{ro}^k denote the representatives of the k^{th} cluster for input (I) $i, i = 1, 2, \dots, S$, and output (O) $r, r = 1, 2, \dots, R$, respectively, where $k = 1$ and 2 , where the first cluster, i.e., $k = 1$, is called a **productive cluster**, while the second cluster, i.e., $k = 2$, is an **unproductive cluster**. Then, using two inputs, TLC and MCD , we can set the representatives of the productive cluster as follows:

$$c_{1I}^1 = TLC_{min}, \quad (13)$$

$$c_{2I}^1 = MCD_{min}. \quad (14)$$

Using two outputs, EDC and CDE , we set

$$c_{1O}^1 = EDC_{max}, \quad (15)$$

$$c_{2O}^1 = CDE_{max}. \quad (16)$$

Now, the representatives of the unproductive cluster for input and output are

$$c_{1I}^2 = TLC_{max}, \quad (17)$$

$$c_{2I}^2 = MCD_{max}, \quad (18)$$

$$c_{1O}^2 = EDC_{min}, \quad (19)$$

$$c_{2O}^2 = CDE_{min}. \quad (20)$$

Now, let α_i and β_r denote the weights assigned to input i and output r , respectively. The more (less) important an input or an output is, the greater (less) α_i or β_r is applied. Then, the weighted Euclidean distances (WEDs) for DMU_j from the representatives of these two clusters are computed as follows:

(i) For input (I),

$$WED_{dI}^k = \sqrt{\sum_{i=1}^2 \alpha_i (x_{id} - c_{il}^k)^2}, k = 1, 2, \quad (21)$$

where x_{1d} and x_{2d} are TLC_d and MCD_d , respectively. If $WED_{dI}^1 < WED_{dI}^2$, DMU_d belongs to the first cluster, C_I^1 . If $WED_{dI}^1 > WED_{dI}^2$, the second cluster, C_I^2 . If $WED_{dI}^1 = WED_{dI}^2$, DMU_j belongs to both clusters.

(ii) For outputs (O),

$$WED_{dO}^k = \sqrt{\sum_{r=1}^2 \beta_r (y_{rd} - c_{ro}^k)^2}, k = 1, 2, \quad (22)$$

where y_{1d} and y_{2d} are EDC_d and CDE_d , respectively. If $WED_{do}^1 < WED_{do}^2$, DMU_d belongs to the first cluster, C_O^1 . If $WED_{do}^1 > WED_{do}^2$, the second cluster, C_O^2 . If $WED_{do}^1 = WED_{do}^2$, DMU_d belongs to both clusters.

We define the total WED ($TWED$) for outputs as the WED between overall maximums and minimums. Then, $TWED$ for the outputs (O) is given by

$$TWED_O = \sqrt{\sum_{r=1}^2 \beta_r (c_{ro}^1 - c_{ro}^2)^2}. \quad (23)$$

Now, WED_{do}^{2R} , the revised WED_{do}^2 for DMU_d , is obtained by subtracting WED_{do}^1 in (22) from $TWED_O$ in (23) as follows:

$$WED_{do}^{2R} = TWED_O - WED_{do}^1. \quad (24)$$

Note that, as WED_{do}^{2R} increases, DMU_d becomes more productive regarding output, since $TWED_O$ in (23) is a constant. Now, we see that, as the ratio of WED_{do}^{2R} to WED_{dl}^1 increases, the more productive DMU_d becomes. The normalized PS, PS_d^N , for DMU_d , can be expressed as

$$PS_d^N = \frac{WED_{do}^{2R}/WED_{dl}^1}{\max_{t \in \{1, 2, \dots, D\}} \{WED_{to}^{2R}/WED_{tl}^1\}}, \quad d \text{ and } t = 1, 2, \dots, D. \quad (25)$$

DMUs are ranked based on the normalized PS, PS_d^N . If the denominator in (25) is too large, the PS for less productive DMUs would be too small. Thus, PS_d^N can be adjusted by increasing WED_{dl}^1 of the most productive DMU, i.e., $PS_d^N = 1$, as long as the ranks generated by the normalized PS do not change.

MACHINE LEARNING ALGORITHM BASED ON PDC (ML-PDC)

As shown in Figure 3, the steps of the training stage based on ML-PDC are as follows.

{Initiation}

Step 1: Set $j = 1$. Set c_{1l}^1 and c_{2l}^1 , c_{ro}^1 and c_{ro}^2 . Set $\mu l = 0$ and $\mu o = 0$. Set $PS_\delta^U = 0$.

{PDC}

Step 2: Set $j = j + 1$. Enter DMU_j in the DMU set and let the DMU set, $w = \{1, 2, \dots, j\}$.

Step 3: Set $\mu l = 0, \mu o = 0$, and $ES_j^C = 0$. If $TLC_j \leq c_{1l}^1$ and/or $MCD_j \leq c_{2l}^1$, set $c_{1l}^1 = TLC_j$, and/or $c_{2l}^1 = MCD_j$. Set $\mu l = 1$. If $EDC_j \geq c_{1o}^1$, and/or $CDE_j \geq c_{2o}^1$, set $c_{1o}^1 = EDC_j$ and/or $c_{2o}^1 = CDE_j$. Set $\mu o = 1$. If $EDC_j \leq c_{1o}^2$, and/or $CDE_j \leq c_{2o}^2$, set $c_{1o}^2 = EDC_j$ and/or $c_{2o}^2 = CDE_j$. Set $\mu o = 1$.

Step 4: If $\mu l \neq 0$ and $\mu o \neq 0$, go to Step 5. Compute $WED_{jl}^1, WED_{jo}^{2R}$, and PS_j^U . If $PS_j^U \geq PS_\delta^U$, replace PS_δ^U with PS_j^U . Compute $PS_j^N = PS_j^U / PS_\delta^U$ to go to Step 6.

Step 5: If $\mu l = 1$ or $\mu o = 1$, compute $WED_{wl}^1, WED_{wo}^{2R}$, and $PS_w^U, \forall w$. Set $PS_\delta^C = \max_{\forall w} \{PS_w^U\}$.

Compute $PS_w^N = PS_w^U / PS_\delta^C, \forall w$.

Step 6: Rank DMU_w based on $PS_w^N, \forall w$. Go to Step 2.

CASE STUDY AND OBSERVATIONS

For a case study, a total of forty-six (46) counties in South Carolina (SC) are grouped based on populations and proximity into twenty (20) counties. For each of the 20 counties, one location is chosen, assuming that all population within the grouped county exists in that location. This paper uses major disaster declaration records in SC. Federal Emergency Management Agency (FEMA) database (FEMA, 2017) exhibits that SC has experienced sixteen (16) major natural disaster declarations, such as hurricanes, floods, tornadoes, etc., from 1964 to 2017. We assume that the county's emergency facility shuts down when a major disaster is declared. Based on the assumption and historical record, each *AFS*'s risk probability (a county or a clustered county) is calculated by dividing the years with major natural disasters by the total years (see Table 1). We select the five potential locations for *CWHs*, based on the proportion of area that each site would potentially cover, population, and proximity to Interstate Highways in SC. In most cases, the numbers of *CWHs* and *RDCs* to be constructed are pre-specified. The *TLC* function given by Eq. (1) is simplified by excluding the fixed cost terms for *CWHs* and *RDCs*. Also, we pre-determine the following parameters for this case study. The maximum numbers of *CWHs* and *RDCs* are set to 2 and 5, respectively. The minimum and maximum *RDCs* a *CWH* must cover are 1 and 10, respectively. Each *RDC* must cover at least 2 ($\ell_j = 2$) and at most 7 ($L_j = 7$) *AFSs*. The capacities of *CWHs* and *RDCs* are set to 2,500K and 1,500K. The maximum effective coverage distance for the case of disasters, ϕ_{cv} in (6), hypothetically set equal to 35 miles to find CDE_{max} .

Table 1: Data for Locations of DRFs

No	City	County	Population (K)	Risk Probability
1	Anderson	Anderson/Oconee/Pickens	373	0.125
2	Beaufort	Beaufort/Jasper	187	0.063
3	Bennettsville	Marlboro/Darlington/ChestDRField	96	0.375
4	Conway	Horry	269	0.375
5	Georgetown	Georgetown/Williamsburg	93	0.438
6	Greenwood	Greenwood/Abbeville	92	0.125
7	Hampton	Hampton/Allendale	33	0.188
8	Lexington	Lexington/Newberry/Saluda	318	0.313
9	McCormick	McCormick/Edgefield	35	0.250
10	Moncks Corner	Berkeley	178	0.313
11	Orangeburg	Orangeburg/Bamberg/Calhoun	123	0.375
12	Rock Hill	York/Chester/Lancaster	321	0.313
13	Spartanburg	Spartanburg/Cherokee/Union	367	0.313
14	Sumter	Sumter/Clarendon/Lee	157	0.375
15	Walterboro	Colleton/Dorchester	135	0.250
16	Aiken†	Aiken/Barnwell	184	0.313
17	Charleston†	Charleston	350	0.250
18	Columbia†	Richland/Fairfield/Kershaw	461	0.375
19	Florence†	Florence/Dillon/Marion	203	0.438
20	Greenville†	Greenville/Laurens	521	0.125

†potential locations for *CWH*

We solve the formulated WGP model for various values of weight, $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$. Each weight changes between 0 and 1 with an increment of 0.1, subject to $\sum_{k=1}^4 \alpha_k = 1$. The ‘Gurobi’ Solver Engine

of Analytic Solver is applied. Two hundred eighty-six (286) configurations arising from the combinations of the setting of α are solved. We can reduce the 286 configurations to sixty-eight (68) consolidated configurations since several cases yield the same values of the four performance measures. Each of the 68 configurations representing DRFs' optimal locations and allocations is considered a DMU.

We select the top 20 DMUs ranked by CE-DEA or ML-PDC, and 23 DMUs are chosen. Table 2 reports each performance measure's value, ES for T-DEA, and cross-efficiency scores (CESs), $CES_{d|All}$, when all 68 DMUs are evaluated (Case 1), and the corresponding rankings, $R_{d|All}$, along with all the results by ML-PDC. As shown in Table 2, T-DEA identifies eight inefficient DMUs whose Ess are less than 1.000. We observe that DMU_{34} , a productive one by T-DEA, is ranked #20, while DMU_{157} , an unproductive one with an ES of 0.9852, is the highest-ranked one as #8 among the inefficient group. CE-DEA identifies DMU_{133} , with the greatest $CES_{d|All}$ of 0.9254, as the top-efficient DMU out of 68 DMUs.

The differences between the rankings generated by CE-DEA and ML-PDC are reported as Range in Table 2. We observe that ML-PDC generates very incompatible rankings for some DMUs, such as $\{DMU_{52}, DMU_{53}, DMU_{81}\}$. To see the robustness of rankings generated by the two methods, we apply both approaches to the top 20 DMUs. We apply CE-DEA to evaluate these top-20 DMUs only (Case 2) and present the CES, $CES_{d|\{20\}}$, and corresponding rankings, $R_{d|\{20\}}$, in Table 3, along with all the results generated by PDC.

For comparison, we report the expected rank, $E[R_{d|All}]$, which is an expected rank based on $R_{d|All}$ in Table 2. The absolute ranking difference, ARD ($=|R_{d|All} - R_{d|\{20\}}|$), is also reported to denote the change in rankings. Table 3 demonstrates CE-DEA issues regarding the consistency of generated rankings with non-zero ARDs. The most notable observation is that the top-ranked DMU has changed from DMU_{133} for Case 1 to DMU_{89} for Case 2. In fact, except for three DMUs, DMU_{42} , DMU_{52} , and DMU_{53} , all DMUs' rankings have changed, demonstrating that the ranks generated by CE-DEA are inconsistent. Surprisingly, DMU_{34} has the largest ARD of 13, and seven DMUs have an ARD of 10 or greater. In contrast, the ML-PDC method generates entirely consistent rankings with all ARDs of zeroes (0). For further investigation, the top five DMUs by both methods in each case are selected (Case 3), and there are eleven (11) DMUs. Table 4 lists the top five DMUs in each Case by CE-DEA and ML-PDC, respectively, showing the robustness of rankings generated by ML-PDC. Note that the decision-makers would not consider the two top-ranked DMUs by the proposed method, $\{DMU_{97}, DMU_{98}\}$, if only CE-DEA is applied to evaluate all the configurations.

Now, we depict the two top-ranked DMUs by both methods, $\{DMU_{89}, DMU_{97}\}$, in Figure 2. Figure 2 shows two sites, {Greenville, Columbia}, are selected as the *CWH* locations by these four DMUs. From the above analysis, there are some interesting managerial implications for decision-makers. DMU_{89} has the highest *CDE* as its advantage, but the lowest *EDC* and the highest *TLC* can be considered a disadvantage. The two DMUs, $\{DMU_{97}, DMU_{98}\}$, consistently ranked #1 and #2 by ML-PDC, show advantages in terms of two inputs, *TLC* and *MCD*, and one output *EDC*, but DMU_{98} with the lowest *CDE* shows its main disadvantage. Each of these top-four *HLNS* configurations has its disadvantages as well as advantages, as described, so the decision-makers should consider them and compare the designs before making the final decision.

Table 2: Top 20 Productive DMUs from 68 DMUs

No	DMU #	TLC(\$)	MCD (miles)	EDC (K)	CDE (K)	DEA Method			ML-PDC		Range
						T-DEA	CE-DEA		PDC		
		Input #1	Input	Output	Output #2	ES	$CES_{d All}$	$R_{d All}$	$PS_{d All}^N$	$R_{d All}^N$	
1	25	393,380	97.3	2482	3092	1.0000	0.8867	14	0.7659	15	1
2	26	393,306	97.3	2482	3092	1.0000	0.8868	13	0.7659	14	1
3	34	400,845	97.3	2482	3092	1.0000	0.8806	20	0.7586	16	4
4	35	398,558	92.5	2520	2852	1.0000	0.8813	18	0.8484	10	8
5	42	363,455	85.6	2723	2185	1.0000	0.8855	15	0.9603	7	8
6	43	345,252	85.6	2755	2139	1.0000	0.9010	7	0.9828	4	3
7	48	366,210	87.0	2525	2476	0.9956	0.8842	16	0.9747	5	11
8	52	424,329	82.7	2151	2258	0.9205	0.7615	49	0.7279	17	32
9	53	427,469	82.7	2156	2258	0.9218	0.7605	50	0.7213	18	32
10	81	331,724	104.0	2811	2827	1.0000	0.9241	2	0.6855	20	18
11	82	376,315	108.0	2957	2846	0.9661	0.8872	12	0.6313	23	11
12	88	380,757	97.3	2451	3092	1.0000	0.8919	10	0.7663	13	3
13	89	343,262	98.6	2647	2889	1.0000	0.9229	3	0.7789	11	8
14	91	341,704	92.5	3009	2038	1.0000	0.8929	9	0.7027	19	10
15	97	330,703	85.6	2720	2139	1.0000	0.9085	4	1.0000	1	3
16	98	335,070	85.6	2734	2139	1.0000	0.9069	5	0.9966	2	3
17	107	355,040	87.0	2451	2476	0.9853	0.8803	21	0.9691	6	15
18	108	357,204	87.0	2513	2476	0.9969	0.8901	11	0.9893	3	8
19	133	329,977	98.6	2585	2889	1.0000	0.9254	1	0.7666	12	11
20	143	301,846	87.1	2637	2094	1.0000	0.9053	6	0.8942	8	2
21	157	299,473	87.1	2568	2094	0.9852	0.8939	8	0.8736	9	1
22	180	316,201	101.0	2904	2226	1.0000	0.8810	19	0.5779	29	10
23	238	308,953	98.6	2536	2530	0.9618	0.8820	17	0.6589	22	5

$CES_{d|All}$: CES with all 68 DMUs evaluated, $R_{d|All}$: Corresponding rank, $PS_{d|All}^N$: PS with all 68 DMUs, $R_{d|All}^N$: Corresponding rank

Table 3: Results for the Top 20 DMUs

No	DMU #	DEA Method					ML-PDC			
		T-DEA	CE-DEA				PDC			
		ES	$CES_{d \{20\}}$	$R_{d \{20\}}$	$E[R_{d All}]$	ARD	$PS_{d \{20\}}^N$	$R_{d \{20\}}^N$	$E[R_{d All}^N]$	ARD
1	25	1.0000	0.9437	5	14	9	0.7659	15	15	0
2	26	1.0000	0.9437	4	13	9	0.7659	14	14	0
3	34	1.0000	0.9397	7	20	13	0.7586	16	16	0
4	35	1.0000	0.9398	6	18	12	0.8484	10	10	0
5	42	1.0000	0.9201	15	15	0	0.9603	7	7	0
6	43	1.0000	0.9277	13	7	6	0.9828	4	4	0
7	48	0.9956	0.9306	11	16	5	0.9747	5	5	0
8	52	0.9206	0.8265	22	22	0	0.7279	17	17	0
9	53	0.9218	0.8261	23	23	0	0.7213	18	18	0
10	81	1.0000	0.9371	8	2	6	0.6855	20	20	0
11	82	0.9662	0.9085	17	12	5	0.6313	22	22	0
12	88	1.0000	0.9458	3	10	7	0.7663	13	13	0
13	89	1.0000	0.9531	1	3	2	0.7789	11	11	0
14	91	1.0000	0.9000	19	9	10	0.7027	19	19	0
15	97	1.0000	0.9308	10	4	6	1.0000	1	1	0
16	98	1.0000	0.9305	12	5	7	0.9966	2	2	0
17	107	0.9854	0.9245	14	21	7	0.9691	6	6	0
18	108	0.9969	0.9339	9	11	2	0.9893	3	3	0
19	133	1.0000	0.9518	2	1	1	0.7666	12	12	0
20	143	1.0000	0.9142	16	6	10	0.8942	8	8	0
21	157	0.9852	0.9033	18	8	10	0.8736	9	9	0
22	180	1.0000	0.8747	21	19	2	0.5779	23	23	0
23	238	0.9661	0.8916	20	17	3	0.6589	21	21	0

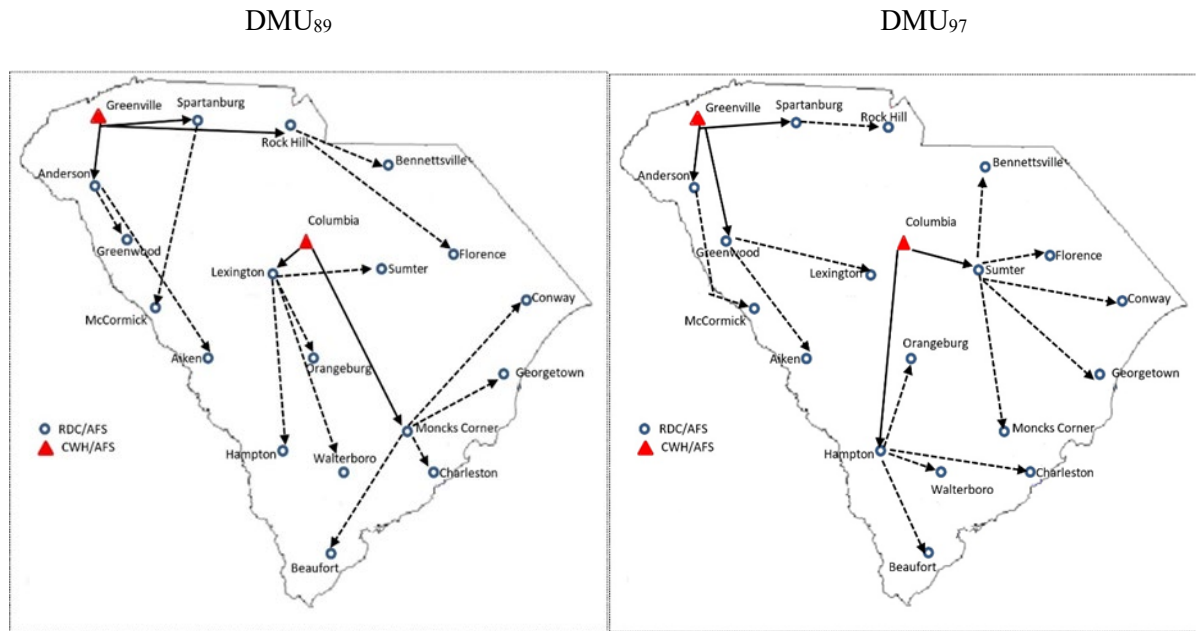
$CES_{d|\{25\}}$: CES with top 20 DMUs evaluated, $R_{d|\{25\}}$: Corresponding rank, ARD (Absolute Rank Difference)= $|R_{d|All} - R_{d|\{25\}}|$

Table 4: Top Five DMUs for Each Case

Rank	CE-DEA			ML-PDC		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	DMU #133	DMU #89	DMU #89	DMU #97	DMU #97	DMU #97
2	DMU #81	DMU #133	DMU #133	DMU #98	DMU #98	DMU #98
3	DMU #89	DMU #88	DMU #88	DMU #108	DMU #108	DMU #108
4	DMU #97	DMU #26	DMU #26	DMU #43	DMU #43	DMU #43
5	DMU #98	DMU #25	DMU #25	DMU #48	DMU #48	DMU #48

Case 1: All 68 DMUs under evaluation, Case 2: Top-20 DMUs under evaluation, Case 3: Top-5 DMUs under evaluation

Figure 2: The six productive humanitarian logistics network system schemes



COMPUTATIONAL EXPERIENCE

We implement the PDC method in an Excel spreadsheet with VBA (Visual Basic for Applications) on an HP Z4 Workstation PC with Intel® Xeon® Gold 5122 and 32GB of RAM installed using a 64-bit version of Windows 10. We randomly generated the values of two inputs and two outputs using a uniform distribution with the minimum and maximum values of all 68 DMUs for the number of DMUs, {68, 100, 200, 300, 400, 500, 700, 1000}. A DEA software using Excel spreadsheets, *DEAFrontier*, is run for the generated DMUs to find efficiency scores on the same computer and compare the computational times between CE-DEA and proposed methods.

The results are listed in Table 6, including the ratio of computational time for CE-DEA to that of PDC. As expected, Table 6 shows that the computational times for PDC are almost negligible compared to CE-

DEA. PDC takes less than seven (7.0) seconds to get the results with one thousand (1,000) DMUs, whereas the DEA software takes more than one hour and five minutes to get the CESs. It also clearly shows that the computational time for CE-DEA sharply increases when the number of DMUs, D , increases. In fact, the computational times seem to grow exponentially in D .

Table 6: Comparison of Computational Time

# of DMUs	(Time unit: second)							
	68	100	200	300	400	500	700	1,000
PDC	0.45	0.66	1.12	1.61	2.12	2.67	3.92	6.80
CE-DEA	45.4	78.5	208.5	421.6	715.6	990.4	2,049.1	3,991.8
Ratio	103.1	118.9	186.2	261.9	337.5	370.9	522.7	587.0

†: Estimated running time

SUMMARY AND CONCLUSIONS

The design of the humanitarian logistics network system (*HLNS*) has become an important strategic decision due to the significant damage inflicted by recent extreme weather-related disasters, including the pandemic. This paper deals with designing resilient, robust, and productive *HLNS* so DRFs can distribute relief items to the affected sites at the right time with the right amount of disaster relief items. Weighted Goal programming (WGP) is applied to generate various network alternatives for planning more balanced *HLNS* schemes. Some authors applied the traditional DEA (T-DEA) method to evaluate these network alternatives. The T-DEA estimates each scheme/DMU on the principle of self-evaluation, ignoring unfavorable inputs/outputs when the efficiency scores are computed.

Consequently, issues related to weak discriminatory power have emerged since multiple DMUs frequently become productive, and these productive DMUs are considered equal regarding efficiency scores. The cross-efficiency DEA (CE-DEA), based on the principle of peer evaluation, was introduced to solve the lack of discriminatory power in T-DEA. But CE-DEA can't eliminate the intrinsic weakness of DEA.

To assess DMUs more consistently without prejudice, this paper proposes an MLA based on the productivity-driven clustering (ML-PDC) method. The proposed method uses the overall minimum and maximum values to represent two clusters rather than the centroid of each cluster. Moreover, the proposed method could evaluate and rank the DMUs more flexibly by allowing the decision-maker to assign unequal weights to each input and output.

Using the actual data available for South Carolina, this paper applies the proposed method and CE-DEA to assess and rank various *HLNS* configurations generated by the WGP model. The proposed ML-PDC exhibits excellent performance over CE-DEA in terms of generating reasonable, consistent, and robust rankings for the DMUs under evaluation. The proposed ML-PDC approach can be valuable for designing *HLNS* and other general supply chain network systems with multiple outputs/inputs.

The limitation of this study comes from the assumption that if a DRF, either *CWH* or *RDC*, is disrupted and can't function, the allocated sites, either *RDCs* or *AFSs*, will not be covered. It implies that if a *CWH* is disrupted, all *RDCs* assigned to this disrupted *CWH* and subsequent *AFSs* will not be covered. The concept of emergency operation should be considered in the case of a DRF shutdown. This paper assumes that only facilities are subject to disruptions, but, in reality, disruption can block the flow of relief items due to route disruptions. Thus, it would be more enhanced research if a disaster backup routing plan is considered for future research.

REFERENCES

References are available upon request from Hong.

Artificial Intelligence-Enhanced Stochastic Optimization to Mitigate Supply Chain Disruptions

Ge Guo

Department of Information Systems and Decision Science

University of Baltimore

11 W. Mt. Royal Ave.

Baltimore, MD 21201

(410)837-5107

cguo@ubalt.edu

Abstract:

The COVID-19 pandemic exposed critical vulnerabilities in global supply chain networks, resulting in unprecedented supply shortages, demand volatility, and logistics disruptions. Managing uncertainty in supply chain networks is a top priority for resilience and competitiveness. This study examines quantitative methodologies to address these evolving challenges. Building on lessons learned during the pandemic, we propose artificial intelligence-driven optimization frameworks. First, a two-stage stochastic mixed-integer optimization model captures uncertainties across supply, demand, production, and logistics. Second, a chance-constrained optimization model focuses on high-probability events, balancing reliability and cost-effectiveness. Finally, a risk-averse optimization model addresses worst-case scenarios to ensure robustness under extreme disruptions. Each model is applied to real-world supply chain test cases of varying scales, assessing computational efficiency and solution quality. The findings offer actionable insights for practitioners aiming to future-proof supply chain networks against not only pandemic-related disruptions but also broader emerging risks in an increasingly interconnected and volatile global economy.

Keywords: Supply chain disruption, stochastic optimization, chance constrained model, risk-averse, artificial intelligence

1. Introduction

The outbreak of the COVID-19 pandemic led to unprecedented supply shortages, demand spikes, and transportation breakdown, simultaneously disrupting supply, demand, and logistics worldwide. Managing supply and demand fluctuations in supply chain networks has never been more crucial. A report in Fortune (Fortune, 2020) indicated that 94% of the companies listed in the Fortune 1000 list were facing supply chain disruptions due to the COVID-19. The major impacts of the COVID-19 pandemic on supply chains include sharp increase of demand for essential goods and healthcare products accompanied by their shortage and declining demand for non-essential products; shortage of material supplies; reduced production capacity and nonavailability of workers; and reduced logistics capacity, loss of physical distribution channels and delays in transportation and distribution (Sawik, 2022). To combat those impacts, manufacturers apply different disruption management strategies, including facility fortification, holding safety stock, direct-to-store delivery, reserved capacity, inventory sharing, and multiple-set covering to mitigate the impact of material flow disruptions and increase resilience of their supply chains (Reza Alikhani, 2021). In addition to the adoption of resilience strategies in supply chain networks, different quantitative methodologies have been proposed to manage the disruptions from the COVID-19 pandemic. Such methodologies include analytical modeling (Paul SK C. P., 2020; Chowdhury M.T.,

2020), simulation modeling (Rahman T, 2021; Ivanov, 2020), optimization modeling (Paul SK C. P., 2021; Paul S.K., 2023; Reza Alikhani, 2021; Sawik, 2022).

The introduction of AI in supply chain management signals a groundbreaking shift, unlocking new levels of efficiency and resilience. It paves the way for innovative approaches to optimizing supply chain operations. AI has been adopted to optimize different areas of supply chain management, including demand forecasting, inventory management, logistics optimization, and supplier relationship management. Nevertheless, mathematical modeling serves as the cornerstone of supply chain optimization. Among a variety of optimization models, stochastic optimization models display advantage in mitigating the pandemic risks. A stochastic model incorporates the uncertainty factors with stochastic parameters that are generated from probability distributions of all possible scenarios. Stochastic programming has been applied to address pandemic disruptions on numerous supply chains (Amin Yazdekhashti, 2021; Mehdi A. Kamran, 2023; Nabil Kenan, 2022) with a variety of uncertainty factors under COVID-19 pandemic considered, such as increasing demand (Amin Yazdekhashti, 2021), supply deficiency (Nabil Kenan, 2022), production capacity (Sawik, 2022), and logistics capacity (Reza Alikhani, 2021).

However, a comprehensive stochastic model that simultaneously addresses all existing uncertainty factors is lacking. More importantly, a systematic comparison among risk-neutral model, risk-averse model, and risk-seeking model of supply chain resilience is needed to help supply chain practitioners to make optimal decisions with different goals or concerns.

To fill the research gap, this study will contribute to the literature as follows:

- A two-stage stochastic mixed integer optimization framework is formulated to integrate all possible disruptions in supply, demand, production, and logistics to optimize the resilience of supply chain networks.
- A risk-averse stochastic model that focuses on worst-case scenarios and a risk-seeking chance constrained optimization model that focuses on optimistic events are proposed and benchmarked against the risk-neutral stochastic model regarding the robustness of their solutions.
- All three proposed optimization models are assessed on real-world supply chain test cases of different scales to evaluate their computational efficiency and solution quality.

2. Literature Review

Since the COVID-19 outbreak in 2019, a considerable number of studies have been performed to protect supply chain networks from pandemic disruptions. The first review article (Priyabrata Chowdhury, 2021) summarized the existing research on the impacts of the COVID-19 pandemic on supply chains and their research methodologies and resilience strategies for managing the impacts of the COVID-19 pandemic. Related literature that incorporated industry 4.0 was reviewed by (Alexander Spieske, 2021) to mitigate risks from COVID-19 pandemic with industry 4.0. The impacts of all types of epidemic outbreaks on supply chains were surveyed in (Maciel M. Queiroz, 2022) and a framework for supply chain management at the times of COVID-19 pandemic was proposed spanning six perspectives, i.e., adaptation, digitalization, preparedness, recovery, ripple effect, and sustainability.

2.1 Supply Chain Management During COVID-19

The COVID-19 pandemic exposed the vulnerabilities of global supply chains, leading to a surge in academic interest around the concept of supply chain resilience. The pandemic caused unprecedented

disruptions across industries. Early studies highlight critical bottlenecks such as labor shortages, factory shutdowns, and transportation constraints (Shih, 2020). For example, the just-in-time (JIT) model, widely adopted for its cost-efficiency, struggled to cope with demand surges and supply shortages, prompting calls for more robust inventory strategies (Ivanov, 2020). The quantitative approaches to manage the disruptions from the COVID-19 pandemics include analytical models (Paul SK C. P., 2020; Chowdhury M.T., 2020), simulation models (Rahman T, 2021; Ivanov, 2020), and optimization models (Paul SK C. P., 2021; Paul S.K., 2023; Sawik, 2022; Reza Alikhani, 2021). Analytical models were proposed with simple strategies to combat impacts of COVID-19 pandemic for essential products such as toilet paper in (Paul SK C. P., 2020) and for food and beverage industry in (Chowdhury M.T., 2020). Simulation techniques were adopted to forecast the impacts of the COVID-19 pandemic on global supply chains of lightning equipment in (Ivanov, 2020) and on essential items of facemasks in (Rahman T, 2021). Deterministic optimization models were applied for recovery planning in manufacturing supply chains for high-demand and essential items such as toilet paper and hand sanitizer (Paul SK C. P., 2021) and for online business operations (Paul S.K., 2023).

2.2 Artificial Intelligence in Supply Chain Optimization

The advent of AI in supply chain management marks a transformative era offering unprecedented opportunities for supply chain resilience and efficiency. Machine learning techniques allow supply chain managers to discover patterns in supply chain data to generate more accurate and reliable forecasts (Deb, 2017; Benbarrad, 2021). Graph algorithms such as minimum cost flow algorithms can optimize the flow of goods and information through supply chain networks while minimizing transportation costs (Vieira, 2019). Game theory can be applied to analyze strategies decisions regarding inventory management, production planning, and capacity allocation in supply chains (Liu, 2022). Evolutionary algorithms can be used to search for near-optimal inventory policies by iteratively adjusting inventory parameters and evaluating their impact on inventory costs (Saracoglu, 2014), and to optimize demand forecasting models by adjusting model parameters and evaluating predictive accuracy (Jalali, 2021). Artificial neural networks is now the most popular computing technology to solve demand forecasting problems in the supply chain optimization (Zhang, 2023) and provide benefits in inventory management, production control, and quality control (Chawla, 2019). Reinforcement learning, an approach to multi-agent systems, has been applied to optimize supply chain operations where agents learn optimal decision-making policies through trial-and-error interactions with their environment and adapt to changing conditions.

2.3 Stochastic Optimization of Supply Chain under Disruptions

Despite the integration of emerging digital technologies such as artificial intelligence, blockchain, and Industry 4.0 analytics, mathematical programming remains at the heart of supply chain optimization. Among all the mathematical modeling, stochastic programming, chance constrained optimization, and robust optimization are chosen to address disruptions in supply chains due to their capability to incorporate uncertainties, with a wide range of applications on retail supply chain (Shafiee Roudbari, 2024; Sawik, 2022), food industry (Rahbari, 2023; Amin Yazdekhasti, 2021), healthcare products (Alizadeh, 2023; Mehdi A. Kamran, 2023), and blood products (Entezari, 2024; Hossein Shirazi, 2021; Nabil Kenan, 2022). The objectives of these stochastic models include minimization of supply chain cost (Song, 2024; Mehdi A. Kamran, 2023; Sawik, 2022), maximization of total profit (Shafiee Roudbari, 2024; Meysam Cheramin, 2021), maximization of customers' satisfaction or minimization of demand shortages (Nabil Kenan, 2022; Entezari, 2024), maximization of service level (Sawik, 2022), minimization of environmental damages (Zeng, 2023; Nayeri, 2021; Shafiee Roudbari, 2024). A variety of uncertainty factors are incorporated such as unpredictable demand (Song, 2024; Zeng, 2023), reduced

supply capacity (Sawik, 2022; Paul S.K., 2023), interrupted transportation (Zeng, 2023; Sawik, 2022), and increased costs (Rahbari, 2023; Vali-Siar, 2022).

2.4 Research Gaps

Despite the extensive body of literature on AI and stochastic optimization in supply chain networks, there is a notable lack of AI-driven optimization models specifically designed to enhance supply chain resilience. Further research is needed to compare various stochastic optimization models for supply chain networks. This study addresses the gap by introducing a risk-neutral two-stage stochastic model that simultaneously tackles challenges such as escalating demand, material shortages, and reduced production and transportation capacities. Additionally, risk-averse and risk-seeking chance-constrained optimization models are developed to cater to the diverse goals and concerns of supply chain decision-makers. AI is integrated into the framework to generate scenarios for uncertain parameters, further enhancing the model's robustness and adaptability.

3. Model Formulation

3.1 Problem Description

Our objective in this paper is to design or redesign a network to minimize the total supply chain costs. Consider a supply chain network with multiple suppliers, manufacturers, distribution centers or warehouses, and retailers or customer zones. The locations of suppliers and customer zones are fixed, while locations of manufacturers and warehouses must be determined from a given set of candidate locations.

Figure 1 visualizes an example of a supply chain network, which consists of a set of nodes and arcs. Suppliers R_1, R_2 ship raw materials or resources to manufacturers M_1, M_2 where raw materials are processed and assembled into products. End products are distributed to warehouses W_1, W_2, W_3, W_4 which replenish the retail stores or customer zones C_1, C_2, C_3, C_4, C_5 . The supply chain network strives to satisfy as much demand as possible from customer zones, as penalty of back order is applied to the demand shortages.

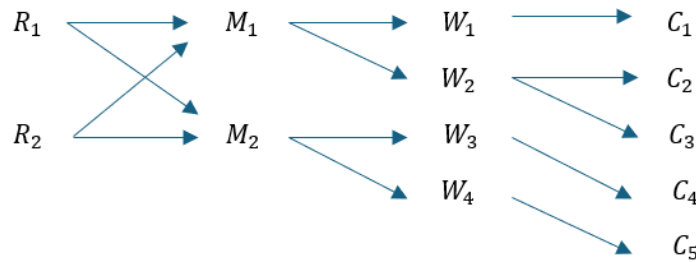


Figure 1. An Example of a Supply Chain Network

We formulate our supply chain optimization problem in a two-stage decision framework. In the pre-disruption phase, we design our supply chain network by making first-stage decisions on facility locations to minimize the fixed costs of opening or renting manufacturing facilities, warehouses or distribution centers, and paths to connect the facilities. In the post-disruption phase, we perform the supply chain planning by making second-stage decisions on the amount of materials to ship from each supplier, the amount of products to produce in each manufacturer, the inventory level at each warehouse, the amount of products to deliver to each customer zone, and the optimal flow decisions at each path. Our second-stage decision are made upon possible realizations of post-disruption scenarios. Thus, our second-stage

model incorporates disruption-related uncertainty factors, including stochastic customer demand, supply volume, production capacity, inventory capacity, and transportation capacity.

Our two-stage optimization framework is formulated into three different models with different risk preferences, risk-neutral stochastic model, risk-averse stochastic model, and risk-seeking chance-constrained model. Risk-neutral models make decisions based solely on expected outcomes, without considering variability. Risk-averse models prioritize minimizing risks, even if it results in worse expected outcomes. Risk-seeking models is driven by the potential for higher rewards in optimistic scenarios and is often associated with more aggressive or speculative decision-making.

The base model of the three different models is formulated as follows:

Sets:

R:	Set of raw material supplier nodes r
M:	Set of candidate manufacturer nodes m
W:	Set of candidate warehouse or distribution center nodes w
C:	Set of customer zone nodes c
A:	Set of possible arcs between node i and j
P:	Set of products p
S:	Set of possible scenarios s

First stage parameters:

FM_m :	fixed cost of production at manufacturer m
FW_w :	fixed cost of opening or renting warehouse w
FA_{ij} :	fixed cost of transporting between i and j
$CR_{r,p}$:	unit purchasing cost of raw materials for product p from supplier r
$CM_{m,p}$:	unit production cost of product p at manufacturer m
$CW_{w,p}$:	unit inventory holding cost of product p at warehouse w
$CA_{ij,p}$:	unit transportation cost of product p shipped on arc (i, j)
$CP_{c,p}$:	unit penalty cost of unsatisfied product p at customer c

Second stage parameters:

$D_{c,p,s}$:	demand of customer c for product p under scenario s
$VR_{r,p,s}$:	supply volume or capacity of supplier r for product p under scenario s

$VM_{m,p,s}$:	production volume or capacity of manufacturer m for product p under scenario s
$VW_{w,p,s}$:	volume or capacity of warehouse w for product p under scenario s
$VA_{ij,p,s}$:	transportation volume or capacity of line from i to j for product p under scenario s
π_s :	probability of scenario s

First stage decision variables:

um_m :	1 if potential location m is selected to open a manufacturing facility; 0 otherwise
uw_w :	1 if potential location w is selected to open a warehouse; 0 otherwise
ua_{ij} :	1 if pathway between i and j is selected to transport the products; 0 otherwise

Second stage decision variables:

$x_{r,p,s}$:	quantity of raw materials for product p to purchase from supplier r under scenario s
$y_{m,p,s}$:	quantity of product p to produce in manufacturer m under scenario s
$z_{w,p,s}$:	quantity of product p to hold at warehouse w under scenario s
$f_{ij,p,s}$:	quantity of product p from i to j under scenario s
$b_{c,p,s}$:	shortage of product p at customer c under scenario s

Constraints:

- Capacity constraints:

- Supplier capacity:

$$x_{r,p,s} \leq VR_{r,p,s}, r \in R, p \in P, s \in S \quad (1)$$

- Production capacity:

$$y_{m,p,s} \leq VM_{m,p,s}um_m, m \in M, p \in P, s \in S \quad (2)$$

- Warehouse capacity:

$$z_{w,p,s} \leq VW_{w,p,s}uw_w, w \in W, p \in P, s \in S \quad (3)$$

- Pathway capacity:

$$f_{ij,p,s} \leq VA_{ij,p,s}ua_{ij}, ij \in A, p \in P, s \in S \quad (4)$$

- Supplier flow from each supplier:

$$\sum_{j \in M} f_{r,j,p,s} = x_{r,p,s}, r \in R, p \in P, s \in S \quad (5)$$

- Production flow conservation:

$$\sum_{i \in R} f_{i,m,p,s} = y_{m,p,s}, m \in M, p \in P, s \in S \quad (6)$$

$$\sum_{j \in W} f_{m,j,p,s} = y_{m,p,s}, m \in M, p \in P, s \in S \quad (7)$$

- Inventory flow conservation:

$$\sum_{i \in M} f_{i,w,p,s} = z_{w,p,s}, w \in W, p \in P, s \in S \quad (8)$$

$$\sum_{j \in C} f_{w,j,p,s} = z_{w,p,s}, w \in W, p \in P, s \in S \quad (9)$$

- Demand constraints:

$$\sum_{i \in W} f_{i,c,p,s} + b_{c,p,s} = D_{c,p,s}, c \in C, p \in P, s \in S \quad (10)$$

- Nonnegativity and binary constraints:

$$um_m, uw_w, ur_r, ua_{ij} \in (0,1) \quad (11)$$

$$x_{m,p,s}, y_{m,p,s}, f_{ij,p,s} \geq 0 \quad (12)$$

Constraints (1)-(4) impose capacity limitations on supply, production, inventory, and transportation flows. Constraints (2)-(4) ensure that only active nodes can function as manufacturers and warehouses, and only active arcs can serve as transportation paths. Constraints (5)-(9) capture the flow equilibrium within the supply chain network. Constraint (10) reflects product backorders or shortages in customer zones. Finally, Constraints (11) and (12) require all first-stage decisions to be binary and all second-stage decisions to be non-negative.

3.2 Risk-Neutral Stochastic Model

The risk-neutral model assigns equal weight to all possible realizations of uncertainty factors, ranging from the worst-case to the most optimistic scenarios. Consequently, its objective function aims to minimize the two-stage expected total supply chain costs across all scenarios $Z_1 + Z_2$, where the first-stage objective cost adds up the fixed costs before the disruption: $Z_1 = \sum_{m \in M} FM_m um_m + \sum_{w \in W} FW_w uw_w + \sum_{i,j \in A} FA_{ij} ua_{ij}$, and the second-stage objective cost is represented by the expected value of post-disruption costs: $Z_2 = \sum_{s \in S} \pi_s \sum_{p \in P} (CR_{r,p} x_{r,p,s} + CM_{m,p} y_{m,p,s} + CW_{w,p} z_{w,p,s} + CA_{ij,p} f_{ij,p,s} + CP_{c,p} b_{c,p,s})$.

3.3 Risk-Averse Stochastic Model

In contrast to the risk-neutral model, the risk-averse model enables decision-makers to prioritize worst-case scenarios, allowing them to make optimal decisions specifically for these adverse conditions rather than considering all scenarios equally. We chose Conditional Value-at-Risk (CvaR) as a coherent risk measure to measure the expected system costs in the worst $100(1 - \alpha)\%$ scenarios for a given probability α . Thus, risk-averse model introduces additional first-stage parameter α to represent the given probability; first-stage variable η to represent the value-at-risk (Var_α), i.e. the minimum α -quantile such that the probability that the second stage cost exceeds it is at least α ; and second-stage parameter v_s , which represents the excess of second stage objective cost of above η .

The risk-averse stochastic model is formulated as follows:

$$\text{Minimize } \eta + \frac{1}{1-\alpha} \sum_{s \in S} \pi_s v_s$$

subject to (1) – (12),

$$\eta + v_s \geq Z_1 + Z_s, s \in S \quad (13)$$

$$Z_1 = \sum_{m \in M} FM_m um_m + \sum_{w \in W} FW_w uw_w + \sum_{i,j \in A} FA_{ij} ua_{ij},$$

$$Z_s = \sum_{p \in P} (CR_{r,p} x_{r,p,s} + CM_{m,p} y_{m,p,s} + CW_{w,p} z_{w,p,s} + CA_{ij,p} f_{ij,p,s} + CP_{c,p} b_{c,p,s}), s \in S$$

3.4 Risk-Seeking Chance Constrained Model

Unlike risk-neutral model or risk-averse model, risk-seeking models such as chance-constrained model prioritize on optimistic or best-case scenarios rather than considering worst-case or average outcomes. Our chance-constrained model requires the demand constraint (10) to be satisfied with a prescribed probability of $\alpha \times 100\%$. In other words, the chance-constrained model aims at optimizing the supply chain cost of the most optimistic $\alpha \times 100\%$ scenarios and does not consider the worst $(1 - \alpha) \times 100\%$ scenarios. In our risk-seeking chance constrained model, we use the same objective function as the risk-neutral model, and replace the demand constraint (10) with the following set of chance constraints:

$$P \left(\sum_{i \in W} f_{i,c,p,s} + b_{c,p,s} \geq D_{c,p,s} \right) \geq \alpha, c \in C, p \in P \quad (14)$$

Assume the net demand is non-negative, the chance constrained (13) can be approximated as follows:

$$\sum_{i \in W} f_{i,c,p,s} + b_{c,p,s} \geq \max(0, D_{c,p,s}) \delta_s, c \in C, p \in P, s \in S \quad (15)$$

$$\sum_{s \in S} (\pi_s \delta_s) \geq \alpha \quad (16)$$

$$\delta_s \in \{0,1\}, s \in S \quad (17)$$

where δ_s is a new binary variable defined by:

$$\delta_s = \begin{cases} 1, & \text{if all constraints in scenario } s \text{ are satisfied} \\ 0, & \text{otherwise} \end{cases}$$

4. Computational Studies

4.1 Scenario Generation

Disruptions in supply chain led to a variety of uncertain factors, including unpredictable demand, disrupted supply, impacted capacities, interrupted transportation, and increasing costs. A stochastic approach is advantageous in optimizing supply chain networks under such uncertainties. Generating accurate and reliable scenarios of uncertain parameters is crucial for optimization of supply chain networks. Traditional forecasting methods, such as time series models, often struggle to capture the complexity of supply chain disruptions. Machine learning, a subset of artificial intelligence, offers a variety of algorithms and techniques that enable learning from historical data, customer demographics, market trends, and other factors to identify patterns, correlations, and trends to make better forecasts. In this study, we use decision trees to generate possible scenarios of demand and capacity for suppliers, manufacturers, warehouses, and transportation.

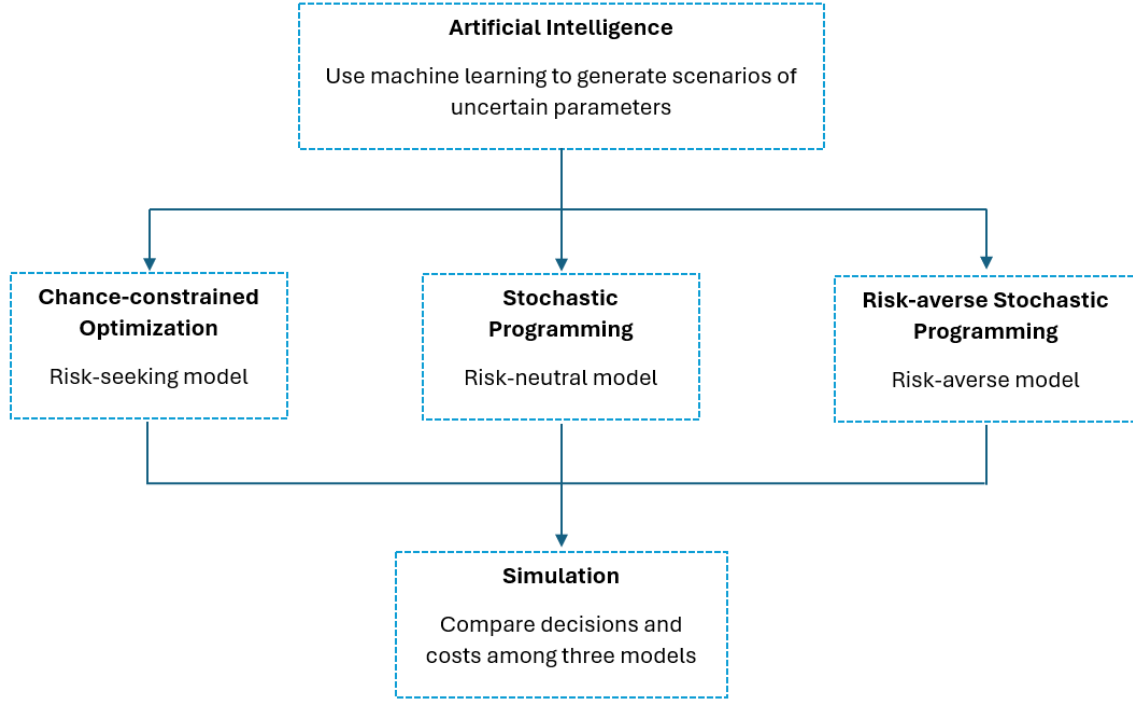


Figure 2. Flow Chart of AI-Enhanced Stochastic Optimization

4.2 Optimal Decisions Across Models

In this section, we compare the decisions and outcomes generated from different optimization models. We consider a supply chain network with two suppliers, (R_1, R_2) , two manufacturers, (M_1, M_2) , four warehouses (W_1, W_2, W_3, W_4) , and 5 customer zones $(C_1, C_2, C_3, C_4, C_5)$, shown in Figure 3. Each customer zone has a demand of 20 units.

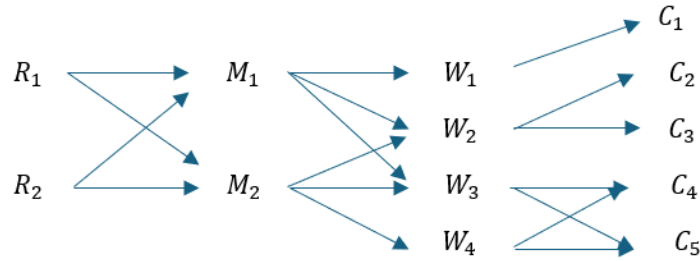


Figure 3: Test Case of Supply Chain Network

To simplify the study, we consider four possible scenarios with varying levels of disruptions:

- Scenario 1: No disruption
- Scenario 2: Disrupted paths from supplier R_1
- Scenario 3: Disrupted paths from supplier R_1 , and decreased production capacity in M_1
- Scenario 4: Disrupted paths from supplier R_1 , decreased production capacity in M_1 , and closure of warehouse W_3

The resulting optimal flow decisions under the worst-case scenario, Scenario 4, are detailed below with the satisfied demand shown in Table 1.

4.2.1 Deterministic Model:

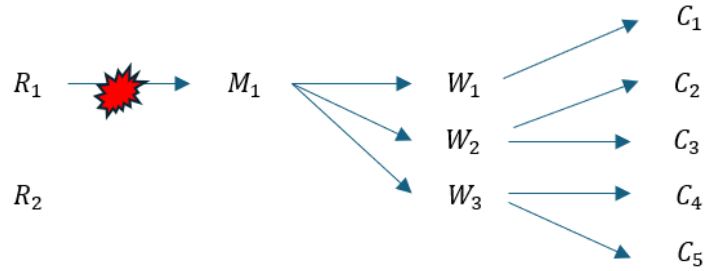


Figure 4(a). Deterministic Decisions in Scenario 4

The deterministic model does not account for any disruptions, basing its decisions solely on Scenario 1. The supply cost from R_1 is cheaper than R_2 , so the deterministic model excludes R_2 from supplying materials. Similarly, the production cost at M_1 is lower than at M_2 , leading the deterministic model to disregard opening the manufacturing facility at location M_2 . As a result, the entire supply chain network becomes reliant on a single path between R_1 and M_1 . In Scenario 4, when this critical link fails, the entire supply chain breaks down, leaving all customer demands unmet.

4.2.2 Risk-Seeking Chance-Constrained Model:

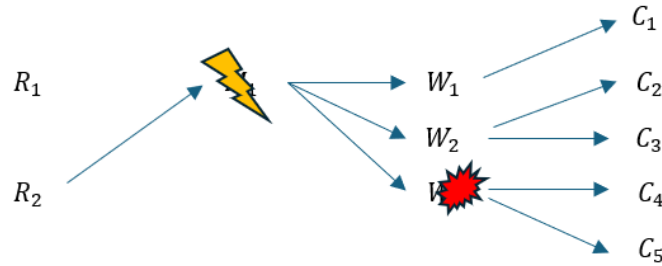


Figure 4(b). Risk-Seeking Decisions in Scenario 4

The risk-seeking chance-constrained model prioritizes the most optimistic scenarios, Scenario 1 and Scenario 2, when making decisions. It relies on R_2 to supply materials because the path from R_1 is disrupted in Scenario 2. All materials from R_2 are directed to M_1 , as the production cost at M_1 is lower than at M_2 . However, when Scenario 4 occurs, M_1 loses half of its production capacity, resulting in unmet demand for half of its customers.

4.2.3 Risk-Neutral Stochastic Model:

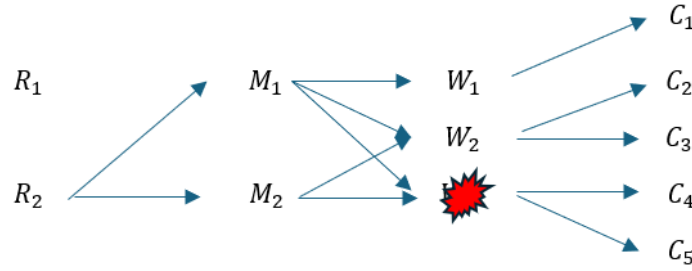


Figure 4I. Risk-Neutral Decisions in Scenario 4

The risk-neutral stochastic model optimizes the supply chain network's average performance by minimizing the expected total costs across all four scenarios. Since the fixed cost of opening a warehouse at W_4 is relatively high and the probability of Scenario 4 is relatively low, the model determines that the optimal decision is to open three warehouses: W_1 , W_2 , and W_3 . However, in Scenario 4, W_3 is closed, preventing customers of C_4 and C_5 , which rely on W_3 , from receiving their required demands.

4.2.4 Risk-Averse Stochastic Model:

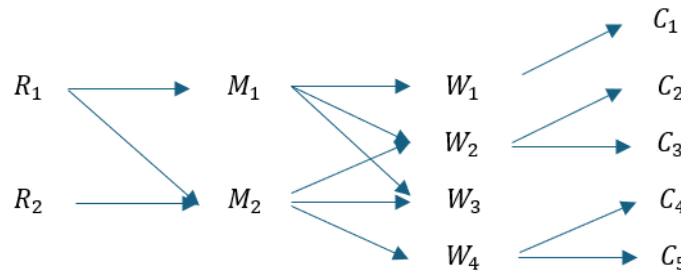


Figure 4(d). Risk-Averse Decisions in Scenario 4

Unlike the models discussed above, the risk-averse model focuses on optimizing performance for the worst-case scenario, Scenario 4. It accounts for all disruptions in this scenario, including the disrupted paths from supplier R_1 , decreased production capacity in M_1 , and closure of warehouse W_3 . Consequently, the risk-averse model is the only one that chooses to open W_4 despite its high opening cost. As a result, the model ensures that all customer demands are fully met.

Table 1 below summarizes the demand satisfaction for each customer zone under the worst-case scenario (Scenario 4). This table highlights the differing approaches of the models: the deterministic model fails to account for uncertainties caused by disruptions; the risk-seeking model prioritizes the most optimistic scenarios; the risk-neutral model focuses on average performance across scenarios; and the risk-averse model aims to ensure demand satisfaction even in the worst-case scenario.

Table 1. Demand Fulfillment in the Worst-Case Scenario Across Different Models

Customer/Model	Deterministic	Risk-seeking	Risk-neutral	Risk-averse
Customer 1	0	20	20	20
Customer 2	0	20	20	20
Customer 3	0	10	20	20

Customer 4	0	0	0	20
Customer 5	0	0	0	20

4.3 Optimal Costs Comparison Across Models

Building on the comparison of flow decisions in the previous section, this section examines the objective costs yielded by different optimization models applied to real-world problems. We evaluate three distinct supply chain industries: retail, agri-food, and anti-epidemic supply chains, each with unique characteristics. The retail supply chain, influenced by the internationalization of suppliers, is particularly vulnerable to supplier shortages. The agri-food supply chain features relatively stable demand for agricultural products but faces significant disruptions in supply chain capacity. In contrast, the anti-epidemic supply chain is driven by sharp demand spikes for products like masks, disinfectants, and protective clothing.

4.3.1 Retail Supply Chain

Based on a real case from Shahrvand Chain Stores in Tehran (Alikhani, 2021), the supply chain network includes two suppliers, two manufacturers, four warehouses, five customer zones, and one product. Uncertainties include supply shortages, reduced manufacturing capacity, and unpredictable demand. We generate 125 scenarios based on five realizations each of supplier capacity, manufacturing capacity, and customer demand.

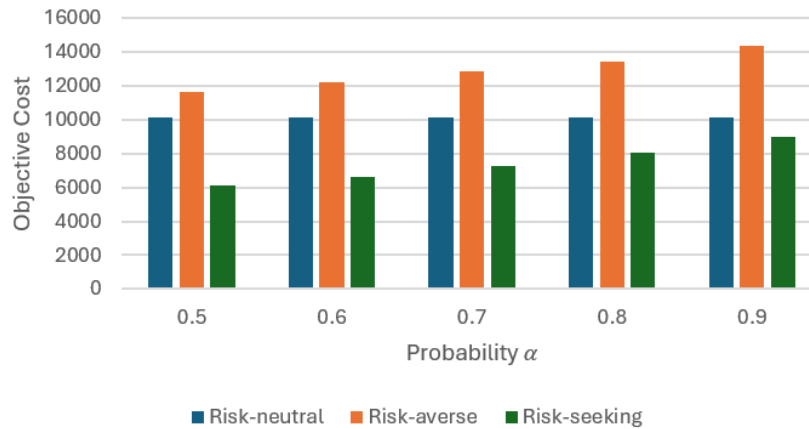


Figure 5(a). Retail Supply Chain Costs Across Models

4.3.2 Agri-Food Supply Chain

Using a canned food supply chain problem in agri-food industry (Rahbari, 2023), uncertainties focus on warehouse capacities. The network includes one canning production unit, six suppliers, three potential warehouses, and fourteen retailers or customer zones. Three products, canned peas, canned green bean, and canned tomato paste are considered in this problem. We generate 25 scenarios based on five realizations each of warehouse capacities.

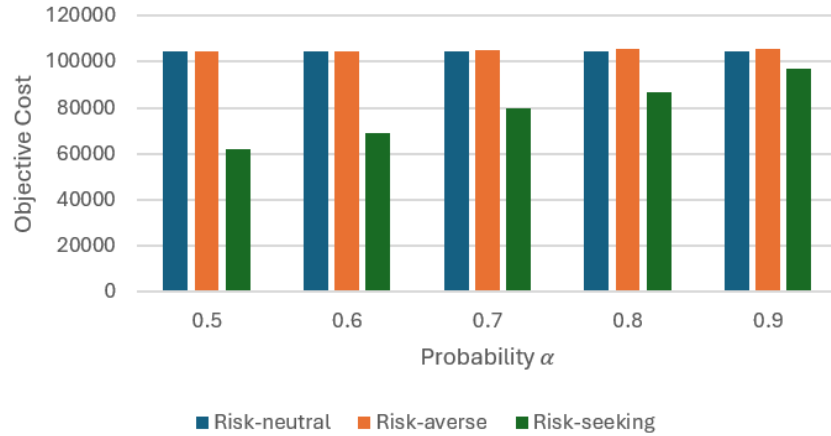


Figure 5(b). Agri-Food Supply Chain Costs Across Models

4.3.3 Anti-Epidemic Supply Chain

A case study in Wuhan, China is adopted as an example of anti-epidemic supply chain problem (Song, 2024). In the actual epidemic prevention process, demand spikes are observed in necessary anti-epidemic materials such as masks, disinfectants, and protective clothing are. Hence, masks, disinfectants, and protective clothing are used as typical anti-epidemic products in this case study. This anti-epidemic supply chain network consists of six suppliers, six distribution centers, and thirteen affected areas. Unlike the previous test cases, manufacturers are not included in this network, as the anti-epidemic materials are supplied directly from the six suppliers. For anti-epidemic supply chain, the largest uncertainty is the demand spikes of anti-epidemic materials after occurrence of epidemic. Five scenarios are generated from the possible demand of customers' demand of anti-epidemic materials.

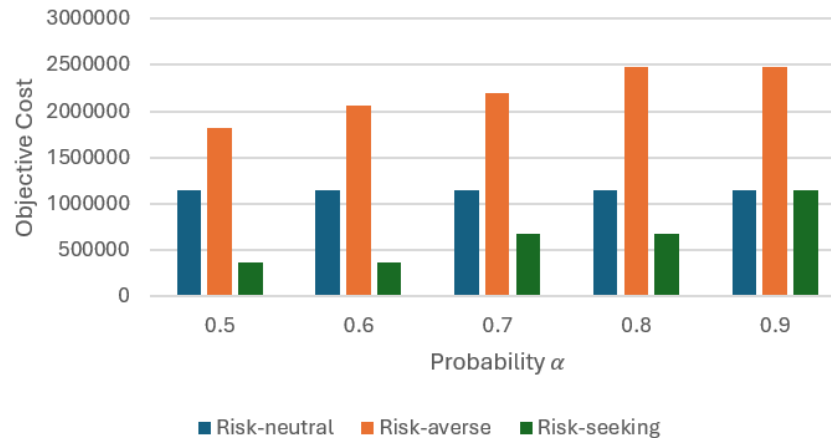


Figure 5I. Anti-Epidemic Supply Chain Costs Across Models

4.3.4 Cost Comparison

As shown in Figures 5(a), (b), and (c), risk-neutral decisions remain constant across varying values of probability α . In contrast, both risk-seeking and risk-averse decisions increase as the probability α rises. Risk-seeking decisions optimize costs for the most optimistic $\alpha \times 100\%$ scenarios, while risk-averse

decisions focus on optimizing costs for the worst $100(1 - \alpha)\%$ scenarios. This distinction explains why risk-seeking models consistently achieve lower objective costs compared to risk-neutral models, whereas risk-averse models result in higher costs than risk-neutral models.

4.4 Run-times Comparison

In addition to the differences in optimal flow decisions, the variations in run-times across different models are also noteworthy. Using three problem settings—retail, agri-food, and anti-epidemic supply chains—we analyzed how run-times scale with the number of scenarios. Computation was terminated if a problem could not be solved within one hour, and such cases are marked as “NA” in the tables below.

As shown in Tables (a), (b), and (c), the run-times for risk-neutral models are consistently the lowest across all three test cases. In contrast, risk-seeking models exhibit significantly higher run-times, particularly for large-scale problems. The increased computational burden of the risk-averse and risk-seeking models can be attributed to the additional constraints and variables introduced in their formulations.

Specifically, the risk-averse stochastic model incorporates additional constraints (13) for every scenario, while the risk-seeking chance-constrained model introduces a binary variable and constraints (15) and (17) for each scenario. Furthermore, the inclusion of constraint (16) in the chance-constrained formulation links scenario-specific decision variables into a single constraint, substantially increasing the computational complexity.

Table 2(a). Run-Times of Retail Supply Chain Optimization

Number of Scenarios	Risk-neutral	Risk-averse	Risk-seeking
125	0.31	0.42	2.49
250	0.46	0.5	7.25
1000	2.61	6.66	NA
2000	3.76	3.9	NA

Table 2(b). Run-Times of Agri-food Supply Chain Optimization

Number of Scenarios	Risk-neutral	Risk-averse	Risk-seeking
25	0.22	0.39	1.62
125	1.38	2.58	74.5
1000	52.3	52.7	NA
2000	38.49	71.56	NA

Table 2I. Run-Times of Anti-epidemic Supply Chain Optimization

Number of Scenarios	Risk-neutral	Risk-averse	Risk-seeking
5	2.4	1.14	5.45
25	11.01	54.55	295.91

125	17.51	NA	NA
1000	48.33	NA	NA

5. Conclusion

This study examines the optimization of supply chain networks under uncertainties caused by disruptions, highlighting the advantages of stochastic modeling techniques. By comparing deterministic, risk-seeking, risk-neutral, and risk-averse optimization models, we demonstrate the varied approaches to addressing uncertainties and their corresponding implications for supply chain performance.

Our findings reveal that deterministic models, while cost-efficient in ideal scenarios, fail to address the complexities introduced by disruptions, leaving customer demands unmet in adverse situations. In contrast, risk-seeking models prioritize optimistic outcomes and fall short in ensuring reliability under worst-case scenarios. Risk-neutral models strike a balance by optimizing average performance, although their reliance on expected costs may result in demand shortfalls in extreme disruptions. The risk-averse model proves to be the most robust in meeting demands even in the worst-case scenario, but leads to significantly higher costs.

The computational studies across retail, agri-food, and anti-epidemic supply chains further highlight the trade-offs between cost and resilience. The varying computational complexities and runtimes of the models underline the practical considerations of choosing an optimization approach tailored to specific industry needs and problem scales. This research underscores the importance of incorporating stochastic models into supply chain optimization, especially in uncertain environments. Future studies can expand on these findings by exploring hybrid models that integrate the strengths of the various approaches or applying advanced AI techniques to enhance scenario generation and decision-making processes.

Bibliography

- Alexander Spieske, H. B. (2021). Improving supply chain resilience through industry 4.0: A systematic literature review under the impressions of the COVID-19 pandemic. *Computers & Industrial Engineering*, 158: 107452.
- Alikhani, R. S. (2021). Retail supply chain network design with concurrent resilience capabilities. *International journal of production economics* , 234: 108042.
- Alizadeh, M. P. (2023). Viable healthcare supply chain network design for a pandemic. *Annals of Operations Research*, 328(1), 35-73.
- Amin Yazdekhashti, J. W. (2021). A multi-period multi-modal stochastic supply chain model under COVID pandemic: A poultry industry case study in Mississippi. *Transportation Research Part E*, 154: 102463.
- Apple falls after iPhone supply-chain disruption drives first revenue miss since 2018.* (n.d.). Retrieved from <https://markets.businessinsider.com/news/stocks/apple-stock-price-iphone-supply-chain-disruption-drives-sales-miss-2021-10>

- Benbarrad, T. S. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, 10(1), 7.
- Chawla, A. S. (2019). Demand forecasting using artificial neural networks—a case study of American retail corporation. *Applications of artificial intelligence techniques in engineering*, pp. 79–89.
- Chowdhury M.T., S. A. (2020). A case study on strategies to deal with the impacts of COVID-19 pandemic in the food and beverage industry. *Operations Management Research*.
- Deb, C. Z. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, 74, pp.902-924.
- Entezari, S. A. (2024). A Bi-objective stochastic blood type supply chain configuration and optimization considering time-dependent routing in post-disaster relief logistics. *Computers & Industrial Engineering*, 188, 109899.
- Fortune. (2020). *94% of the Fortune 1000 are seeing coronavirus supply chain disruptions: Report*. Fortune.
- Hossein Shirazi, R. K. (2021). A stochastic bi-objective simulation–optimization model for plasma supply chain in case of COVID-19 outbreak. *Applied Soft Computing*, 112: 107725.
- Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: a simulation-based analysis on the coronavirus outbreak (COVID-19/SARSCov-2) case. *Transportation Research Part E*, 136:101–922.
- Jalali, S. A.-k. (2021). A novel evolutionary-based deep convolutional neural network model for intelligent load forecasting. *IEEE Transactions on Industrial Informatics*, 17(12), pp.8243-8253.
- Liu, P. H. (2022). Cost reduction of inventory-production-system in multi-echelon supply chain using game theory and fuzzy demand forecasting. *International journal of fuzzy systems*, 24(4), pp.1793-1813.
- Maciel M. Queiroz, D. I. (2022). Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review. *Annals of Operations Research*, 319:1159–1196.
- Mehdi A. Kamran, R. K. (2023). A new vaccine supply chain network under COVID-19 conditions considering system dynamic: Artificial intelligence algorithms. *Socio-Economic Planning Sciences*, 85: 101378.
- Meysam Cheramin, A. K. (2021). Resilient NdFeB magnet recycling under the impacts of COVID-19 pandemic: Stochastic programming and Benders decomposition. *Transportation Research Part E*, 155: 102505.
- Mohammad Mahdi Vali-Siar, E. R. (2022). Sustainable, resilient and responsive mixed supply chain network design under hybrid uncertainty with considering COVID-19 pandemic disruption. *Sustainable Production and Consumption*, 30: 278–300.
- Nabil Kenan, A. D. (2022). The supply chain of blood products in the wake of the COVID-19 pandemic: Appointment scheduling and other restrictions. *Transportation Research Part E*, 159: 102576.

- Nayeri, S. T. (2021). A multi-objective fuzzy robust stochastic model for designing a sustainable-resilient-responsive supply chain network. *Journal of Cleaner Production*, 311, 127691.
- Paul S.K., M. M.-S. (2023). A recovery plan- ning model for online business operations under the COVID-19 outbreak. *International Journal of Production Research*, 61(8): 2613-2635.
- Paul SK, C. P. (2020). Strategies for managing the impacts of disruptions during COVID-19: an example of toilet paper. *Global Journal of Flexible Systems Management*, 21(3): 283-293.
- Paul SK, C. P. (2021). A production recovery plan in manufacturing supply chains for a high-demand item during COVID-19. *International Journal of Physical Distribution and Logistics Management*, 51(2): 104–25 .
- Priyabrata Chowdhury, S. K. (2021). COVID-19 pandemic related supply chain studies: A systematic review. *Transportation Research Part E*, 148: 102271.
- Rahbari, M. K. (2023). Robust optimization and strategic analysis for agri-food supply chain under pandemic crisis: Case study from an emerging economy. . *Expert Systems with Applications*, 225, 120081.
- Rahman T, T. F. (2021). An agent-based model for supply chain recovery in the wake of the COVID-19 pandemic. *Computers and Industrial Engineering*, 158:107401.
- Reza Alikhani, S. T. (2021). Retail supply chain network design with concurrent resilience capabilities. *International Journal of Production Economics*, 234.
- Saracoglu, I. T. (2014). A genetic algorithm approach for multi-product multi-period continuous review inventory models. . *Expert Systems with Applications*, 41(18), pp.8189-8202.
- Sawik, T. (2022). Stochastic optimization of supply chain resilience under ripple effect: A COVID-19 pandemic related study. *Omega*, 109.
- Shafiee Roudbari, E. F. (2024). Designing a multi-objective closed-loop supply chain: a two-stage stochastic programming, method applied to the garment industry in Montréal, Canada. *Environment, Development and Sustainability*, 26(3), 6131-6162.
- Shih, W. C. (2020). Global supply chains in a post-pandemic world. *Harvard Business Review*, 98(5), 82-89.
- Song, M. Y. (2024). Robust optimization model of anti-epidemic supply chain under technological innovation: learning from COVID-19. *Annals of Operations Research*, 335(3), 1331-1361.
- Vali-Siar, M. M. (2022). Sustainable, resilient and responsive mixed supply chain network design under hybrid uncertainty with considering COVID-19 pandemic disruption. *Sustainable production and consumption*, , 30, 278-300.
- Vieira, C. L. (2019). Minimum-cost flow algorithms: a performance evaluation using the Brazilian road network. *World Review of Intermodal Transportation Research*, 8(1), pp.3-21.
- Where is the iPhone made?* (n.d.). Retrieved from <https://www.lifewire.com/where-is-the-iphone-made-1999503>

- Zeng, L. L. (2023). Designing a resilient and green coal supply chain network under facility disruption and demand volatility. *Computers & Industrial Engineering*, 183, 109476.
- Zhang, Z. S. (2023). An intelligent framework for analyzing supply chain resilience of firms in China: a hybrid multicriteria approach. *The International Journal of Logistics Management*, Vol. 34 No. 2, pp. 443-472.

Abstracts



Bitcoin and Lottery Stocks – Competition for The Investor’s Attention

Tao Li¹, Wenxiu Nan²

¹SUNY New Paltz, New Paltz, NY, USA. ²Salisbury University, Salisbury, MD, USA

Abstract

We examine the attention allocation between two speculative investments: Bitcoin and lottery-like stocks. In the traditional stock market, speculators prefer lottery-like stocks as they have a small chance of generating extreme profits. Meanwhile, Bitcoin from the digital financial market exhibits similar return patterns, which also makes it a good choice for speculation. We hypothesize a substitute effect that extreme Bitcoin returns as attention shocks compete for retail investors’ attention to lottery-like stocks.

Following the literature on “categorical learning”, we use the increase of stock return comovement to proxy for change of attention. Based on dynamic conditional correlation (DCC) models, we assess the abnormal stock return comovements around Bitcoin shocks. We employ five measures for the stock lottery-like features: the maximum daily return in the most recent month, price level, idiosyncratic volatility, idiosyncratic skewness, and expectation on jackpot returns. Our findings yield mixed conclusions with different lottery-like measures. Based on ex-ante measures, i.e., idiosyncratic volatility, idiosyncratic skewness, and the expected probability of jackpot returns, the evidence supports our hypothesis that Bitcoin shocks attract more attention away from lottery-like stocks than other stocks. However, the ex-post lottery-like measures such as maximum daily returns and price levels produce the opposite outcomes. We attribute these mixed findings to the timeliness of firm-specific information being priced into the lottery-like features. Our further explorations through the earnings announcement, abnormal trading activities, and gains and losses of the Bitcoin shocks produce similar empirical evidence that supports the substitute effects of Bitcoin on lottery stocks, as well as the sensitivity of the choice of the lottery-like measures on the conclusion.

Being the first study to examine the attention competition between Bitcoin and lottery-like stocks, our findings contribute to both the research stream of the empirical investigation of investors’ rational inattention behaviors and the studies related to lottery-like stocks.

The Spillover Effect of Dual Holdings on Analysts' Information Environment

Seung Won Lee¹, Heeick Choi², Sunghoon Joo³, Hakjoon Song³

¹Penn State Harrisburg, Middletown, PA, USA. ²Sungkyunkwan University, Seoul, Korea, Republic of. ³California State University, Dominguez Hills, Carson, California, USA

Abstract

This study explores the effect of institutional dual ownership on analysts' information environment. We find that dual ownership enhances the accuracy of analysts' forecasts and decreases their dispersion. Both propensity score-matched sample and difference-in-differences tests using financial institution mergers as quasi-natural experiments confirm our findings. Furthermore, cross-sectional analyses suggest that the positive effects of dual ownership on forecast quality are more pronounced in firms with severe information asymmetry, weak external monitoring, and high conflicts of interest between equity and debt holders. Additionally, path analysis reveals that improved accounting quality and management guidance serve as mediating factors that partly explain the relationship between dual ownership and analysts' forecast quality. Overall, our results suggest that dual ownership enriches the information environment for analysts by promoting more efficient monitoring and better voluntary disclosure, consequently improving the quality of analysts' earnings forecasts.

From Investigation to Incarceration: Beyond A Single Agency Prosecuting
Financial Institution Related Crimes

Renee Flasher, Olivia Bolinsky

Penn State Harrisburg, Middletown, PA, USA

Abstract

This paper examines the landscape for multi-entity investigations for the prosecution of financial institution related crimes. We leverage data from the Office of the Special Inspector General for the Troubled Asset Relief Program (SIGTARP) to create a dataset of 88 prosecutions resulting from frauds arising from institutions accepting funds as part of the 2008 financial bailout of financial institutions. Our results highlight the complexity and reality of interagency cooperation required to prosecute these types of crimes. This provides a potential explanation why few criminals are held to account as one entity alone does not typically complete and control the investigation.

Economic Determinants of profitability of listed banks in the United States: A Panel Data Approach

Ikechukwu Ndu¹, Emmanuel Anoruo², Chiaku Chukwuogor³

¹William Paterson University, Wayne, New Jersey, USA. ²Coppin State University, Baltimore, Maryland, USA. ³Eastern Connecticut State University, Willimantic, Connecticut, USA

Abstract

Unlike previous studies to date, this research paper investigates the relationship between financial performance in terms of profitability and the economic determinants for listed banks in the United States since the end of 2014 when banks are required to maintain minimum capital ratios comprising a common equity tier 1 ratio of 4 percent, a tier 1 ratio of 5.5 percent, a total ratio of 8 percent, and a leverage ratio of 4 percent.

Over the past year, the U.S. Federal Reserve has hiked interest rates to combat high inflation. Increased rates mean higher borrowing costs, which puts a strain on companies, especially venture-backed start-ups that need funds from banks. Increased interest rates reduced the value of the bonds in which banks invested their customers' deposits when rates were lower. These factors contributed to the collapse of Silicon Valley Bank, Signature Bank and First Republic Bank. The failure of these high-profile banks caused volatile stock markets and widespread panic, especially at regional banks where institutional customers had large amounts of uninsured deposits. This study is important because of the relative paucity of research on this topic. The results of this study will therefore be useful to academics and assist governments, central banks, standard setters, regulators, supervisors and industry practitioners with formulating policy.

The methodology implements a panel data approach. Data are sourced from published annual financial statements of sampled U.S. listed banks using banking research databases for the period 2015 to 2022 and quoted on the NASDAQ and NYSE stock exchanges.

Buy-Now-Pay-Later (BNPL) Lending: Innovation and Opportunity

Subramanian Narayanaswamy

New Jersey Institute of Technology, Newark, NJ, US

Abstract

Buy Now Pay Later (BNPL) is a more recent lending innovation driven by advancements in Financial Technology (FinTech). Since the 2015s, BNPL has witnessed meteoric growth and has come to cannibalize traditional payment methods such as credit cards. BNPL offers a refreshed product concept adopted mainly by younger millennials and Gen Z. It is positioned more as a lifestyle product than a serious credit. It is easy to use and caters to the instant gratification needs of consumers. BNPL is a payment product, not a credit product; hence, it is relatively unregulated. The companies offering BNPL work with merchants and consumers (tripartite relationship); therefore, it poses more operational complexity than a traditional credit product. This study reviews three vital functional aspects of BNPL (Consumer Behavior, Regulation and Consumer Protection, and Soft Inquiries and Credit Bureau Reporting) and proposes directions for further research. Despite its success, there is limited literature available on BNPL, and it is understudied. To the best of my knowledge, this is the first comprehensive literature review that examines various functional aspects of BNPL lending innovation, integrating practical insights with academic research. BNPL is the fastest-growing payment product and has a huge societal impact. A thorough review is essential to encourage further innovation and promote responsible credit practices.

Exploring Corporate Capital Structure and Overleveraging in the Pharmaceutical Industry

Samar Issa¹, Hussein Issa²

¹Saint Peter's University, Jersey City, NJ, USA. ²Rutgers University, New Brunswick, NJ, USA

Abstract

This paper develops an empirical model of corporate capital structure, optimal debt, and overleveraging, covering the time span 2000 to 2018 across the pharmaceutical industry. Estimated using a sample of twenty firms, the model sheds light on an industry-specific default risk, measuring overleveraging as the difference between actual and optimal debt. The calculated corporate excess debt has largely been moving up, spiking around the global financial crisis in 2008 and starting to recover afterwards, but started to have a risky upward trend more recently. The analysis presented in this paper reveals a concerning trend in the pharmaceutical industry, with corporate excess debt steadily increasing over the past two decades, particularly peaking during the 2008 crisis and after 2013. This suggests that excess debt, rather than its mere presence, played a significant role in the financial meltdown of 2007–2009 and the industry overall is vulnerable. These findings have broader policy implications for future macroeconomic scenario assessments, highlighting the importance of understanding industry-specific default risk and optimal debt levels.

A Leveraged Bootstrap Causality approach to investigate relationships between the Consumer inflation, stock market and real estate market during March 2020 – Oct 2024 which cover the rise of the inflation and the decay.

Howard Lee

Commonwealth University of PA, Bloomsburg, PA, USA

Abstract

The main purpose of this study is to examine the causal relationships among the behaviors of 12-month percentage change in Consumer Price Index (CPI), S&P500 and REIT. The relationships were investigated using the Granger causality test with a leveraged bootstrap approach developed by Hacker and Hatemi-J. The results indicate that at the rise period of inflation, the performance of the CPI is driven by the movements of the S&P500 highly. After the first year of the rise of the CPI, the stock market lost some influence as the CPI took the place of the leading factor of the stock market and the real estate market. Our findings indicate that the first year of the declaration of the Covid-19 Pandemic which is the start point of the rise of all of the inflation, stock market and real estate market; the stock market has great influence on the CPI and the real estate market. As time goes by, both the stock market and the REIT were driven by the CPI greatly. This research provides perspective for building investment portfolios after a global event occurs.

A Reexamination of Foreign Exchange Hedging by Multinational Enterprises
Augustine Arize¹, John malindretos², Theodoros Stamatopoulos³, Chuanqian
Zhang²

¹Texas A&M, Commerce, Texas, USA. ²William Paterson U, Wayne, NJ,
USA. ³University of West Attica, Athens, Greece

Abstract

This is a survey of hedging by US multinational companies.

Predictability of Financial Restatements: Conventional vs. Machine Learning

Khaled Abdou

Penn State University – Berks, Reading, PA, USA

Abstract

This paper examines various models for predicting financial restatements. It compares conventional econometric techniques, such as Logit analysis, with deep learning methods such as neural networks, Sector Vector Machine, and Bootstrap Forest. The paper empirical analysis uses restatement data from Audit Analytics, Compustat and IBES for the period between 2006 and 2023. Preliminary results suggest that not all machine learning models outperform the conventional methods. For instance, the deep learning neural network performed the best.

Enhanced Probability of Default Estimation with Esscher Transform and EM Algorithm

Sung Ik Kim

Louisiana State University Shreveport, Shreveport, LA, USA

Abstract

This research aims to enhance the accuracy of Probability of Default (PD) estimation models, addressing critical needs in today's complex financial environment. Traditional PD models rely on assumptions that often misrepresent real market data, risking inaccurate predictions and ineffective risk management. This project introduces a refined approach, integrating the Esscher transform and Expectation-Maximization (EM) algorithm within firm value models, to capture non-Gaussian data patterns and dynamic correlations. By improving default prediction accuracy, these methods contribute to economic stability, equipping financial institutions with resilient tools to assess credit risk, manage market volatility, and prevent systemic disruptions during financial crises.

A New Data Envelopment Analysis Approach that Ranks Mutual funds and Mutual Fund Companies

Richard Muszynski III

Wilkes University, Wilkes-Barre, PA, USA

Abstract

Mutual funds and the companies that provide these investments to investors have been one of best options for investment growth. Mutual fund popularity has increased due to the diverse selection of stocks, bonds, and other securities mainly in part because of the reduced investment risk while maintaining an acceptable overall return. Yet, selecting a mutual fund that matches an investors' beliefs and expectations can be complicated. We implement a new data envelopment analysis (DEA) approach that consists of solving two nonparametric linear programs to determine optimal mutual funds for investors to invest. A ranking is developed by calculating and sorting efficiency scores in order to assess mutual funds across several mutual fund companies by solving the input-oriented DEA model (Banker et al., 1984) and the output-oriented DEA model (Banker et al., 1984). After ranking the set of mutual funds for each mutual fund company, we then rank mutual fund companies by taking averages, medians, and standard deviations of the efficiency scores of individual mutual funds of each company to determine overall and consistent mutual fund company performance. We conduct several iterations for various input and output variables in order to determine the overall average, median, and standard deviation of many potential variable combinations, and we determine the companies that offer the most consistent performing mutual funds and ones that are highly regarded, i.e., the mutual funds that produce the highest average rate of return.

Ethics of ChatGPT in CPE qualified assessment for CPAs

Dov Fischer

Brooklyn College, Brooklyn, NY, USA

Abstract

AI presents an ethical challenge to the CPA practitioner taking CPE. As accounting rules continue to evolve and become more challenging to the CPE student, there is temptation to use AI for qualified assessment. The temptation is compounded by the silence of NASBA (2024) on the subject. In the most recent standards revised in January 2024, there is no mention of AI. Any prohibition on the part of the student against using AI for qualified assessment is at most implicit rather than explicit.

Beyond Linearity: Modeling Key U.S. Economic Indicators with Rational Curves

Taoufik Meklachi, Brian Adams

Pennsylvania State University Harrisburg, Middletown, PA, USA

Abstract

In this study, we present an efficient and highly accurate method for fitting discrete data using rational functions. The core novelty of the approach lies in scaling positive explanatory data to the interval $[0,1]$, thereby reformulating the problem as a linear homogeneous system. After deriving the fitting function, a rescaling process restores the data to its original scale. Solutions to the problem are obtained from the augmented matrix associated with the homogeneous equation, with each solution of a predefined degree yielding a rational fitting function. The optimal function is selected through a pointwise error analysis. This approach allows for exploration across a wide range of rational functions with varying degrees, ensuring flexibility and precision in finding the best fit. Furthermore, any linear combination of solutions within the null space basis generates new fitting functions, offering adaptability to meet specific constraints. We demonstrate the effectiveness of this method by applying it to U.S. economic metrics, including GDP, unemployment, and inflation rates.

Understanding the Formation of Joint Ventures: ASU 2023-05

Nathan Slavin¹, JIANING FANG²

¹HOFSTRA UNIVERSITY, HEMPSTEAD, NEW YORK, USA. ²KEAN UNIVERSITY, NEWARK, NJ, USA

Abstract

The Financial Accounting Standards Board (FASB) has issued a new standard on accounting for the formation of joint ventures (JV), Accounting Standards Update (ASU) No. 2023-05, *Business Combinations—Joint Venture Formations, (Subtopic 805-60), Recognition and Initial Measurement*. Prior to the ASU, there were no uniform standards. The joint venture could record the assets received and the liabilities assumed at the carrying values reported by the venturers or at the fair market values. Where the JV used carrying values, additional complexity was created when the JV recorded any income using the equity method. The JV would adjust the income effects caused by the difference between the fair values and carrying values of the JV's net assets. As of January 1, 2025, all JV formations are required to record all assets received and liabilities assumed at their respective fair values.

Our presentation will provide a comprehensive analysis of the new accounting for joint venture formations. Joint ventures must now be a new business entity. On the JV formation date, any excess fair value of the JV over its identifiable assets received and liabilities assumed must be recognized as goodwill. We will provide an illustrative example for JV formations with and without a noncontrolling interest (NCI). All necessary journal entries for both the JV and its venturers will be discussed.

Asset Impairment Created by Climate Change: New Challenges for Recording and Reporting

U. J. Stagliano

Saint Joseph's University, Philadelphia, PA, USA

Abstract

Change, even with the best intentions, can lead to unintended consequences. This is undoubtedly the outcome as the world economy moves to shedding dependence on fossil fuels for production of power, products, and services.

Altering output focus has been common since the beginning of the Industrial Revolution. Now, the paradigm shift occurring with movement to a net-zero carbon macroeconomic regime will “strand” trillions of dollars of productive resources. Beyond in-place assets that likely will lose their usefulness, even greater amounts of value will disappear when unrecovered fossil-fuel reserves are abandoned. Even if consumers do not adjust their behavior and acquisition habits, the upstream creation of products/services is undeniably moving to new processes that will leave a vast array of unusable resources in its wake.

What are stranded assets? These are physical properties and intangible rights that have suffered premature write-downs and devaluations from causes not anticipated when they originally were acquired. In some cases, the previously valuable asset actually has become a liability through no fault or action of its owner. Stranding may occur because of environmental challenges, government edict, legal issues, or changes in social norms.

How will we accommodate the obsoleting of perfectly useful property, plant, and equipment in the name of essential progress toward a more environmentally friendly means of providing the goods and services demanded by the world's consumers? More to the point of business management, what will be the financial impact of dealing with write-offs, impairment charges, reduced net cash flows, loss of surety value, and reputational damages that ensue?

This project explores the broad question of accounting and reporting for the now-expected stranding of assets in the energy generation, public utility, and fossil-fuel resources industries. The research contributes to the on-going debate regarding the intersection of profit generation and sustainable development.

Cross-Border Accounting Scandals and Frauds
Ikechukwu Ndu

William Paterson University, Wayne, New Jersey, USA

Abstract

This research paper investigates the underlying reasons and causes of the continued persistent perpetration of cross-border accounting scandals and frauds by managers of the firms to date. A multi-case analysis methodological approach is adopted. The objective is to understand the accounting situation of the cross-border accounting scandal and fraud cases in the context of its total historical experience. The findings will be of interest to regulators, academics, practitioners, and students who seek to increase the understanding of accounting's past cross-border accounting scandals and implement that understanding to explain accounting's present and its possible future development.

Similarities between the development of the Jewish calendar and generally
accepted accounting principles (GAAP)

Dov Fischer

Brooklyn College, Brooklyn, NY, USA

Abstract

Modern financial statements evolved over the past 600 years from basic bookkeeping techniques into intricate methods of accrual accounting. Jewish calendar reckoning had a similar evolutionary period from approximately 360 CE to 960 CE. While the two systems differ fundamentally, there are several similarities that can help us better understand the nature of both systems. These similarities include gradual development; a historic mist of early origins; the reconciliation of multiple cycles; the measurement of states and flows; and inferior status of the two disciplines to their respective legal realms. The history of the Jewish calendar can shed light on future development of accounting systems.

Empowering Learning Business Analytics through ChatGPT

Rose Sebastianelli, Nabil Tamimi, Iordanis Petsas

University of Scranton, Scranton, Pa, USA

Abstract

Generative artificial intelligence (GenAI), epitomized by ChatGPT, has introduced significant opportunities and challenges in higher education, particularly in business analytics instruction. While concerns about academic integrity and cognitive dependency persist, ChatGPT's ability to translate natural language into code positions it as a transformative educational tool. This study explores the potential of ChatGPT to enhance learning in business analytics by complementing traditional software like Excel, Oracle Crystal Ball, and Minitab. A faculty-guided, intentional integration approach is proposed, emphasizing foundational understanding before leveraging ChatGPT for advanced analytics tasks.

By conducting structured in-class exercises, students compare outputs from ChatGPT with those generated by traditional software across various analytics domains, including supervised and unsupervised machine learning. Examples include logistic regression, simulation, time series forecasting, k-means clustering, and market basket analysis. Findings highlight ChatGPT's ability to democratize access to Python and R functionalities, offer additional analytical insights, and encourage interactive learning through iterative prompting. However, limitations include occasional inaccuracies, dependency risks, and reduced student engagement with fundamental concepts.

The study underscores the importance of faculty intervention to mitigate over-reliance on GenAI. Through guided comparisons, students develop AI literacy while critically evaluating the accuracy, interpretability, and usability of ChatGPT outputs. This pedagogical strategy aligns with the growing demand for analytics talent by fostering deeper learning and preparing students for real-world applications.

This research concludes that integrating ChatGPT into business analytics education can empower learning while addressing challenges inherent in GenAI adoption. Faculty-led frameworks that balance traditional methods with innovative AI tools are essential to enhance student engagement and equip learners with the skills to navigate an evolving technological landscape.

Advancing Stock Price Prediction: A Hybrid LLM and Transformer-LSTM Framework with Multi-Task Learning

Rashedul Hasan, Asil Oztekin

University of Massachusetts Lowell, Lowell, MA, USA

Abstract

Stock price prediction has long been a critical area of research, with recent advancements leveraging social media sentiment to forecast price movements. Despite significant progress, existing approaches face challenges in sentiment classification accuracy and capturing temporal dependencies in stock prices. This paper proposes a novel framework that addresses these limitations by integrating Large Language Models (LLMs), such as GPT and FinBERT, with hybrid Transformer-LSTM architectures. By incorporating enriched contextual sentiment features extracted from LLMs and multi-modal data, including technical indicators, the proposed method offers a more robust mechanism for predicting stock price movements.

The framework introduces a multi-task learning approach to simultaneously optimize sentiment classification and stock prediction tasks. Extensive experiments conducted on Tesla stock data demonstrate notable improvements over existing methods, with significant gains in sentiment classification accuracy (F1 score) and stock price prediction metrics (MSE and RMSE). Furthermore, the model's ability to adapt to complex, non-linear patterns in social media sentiment and financial data highlights its scalability and practical applicability.

This research sets a new benchmark for leveraging advanced LLMs and hybrid neural architectures in stock market prediction, offering a pathway for more accurate and reliable financial decision-making tools. The findings underscore the potential of multi-task learning and multi-modal data fusion in transforming market sentiment analysis and stock price forecasting.

Advancing E-Waste Management and Fraud Prevention with Ensemble Machine Learning in Remanufacturing

Elif Kongar¹, Gazi Duman¹, Surendra Gupta²

¹University of New Haven, West Haven, CT, USA. ²Northeastern University, Boston, MA, USA

Abstract

As global electronic waste (e-waste) production is projected to exceed 82 million tons annually by 2030, the need for innovative, sustainable management strategies has never been more urgent. Remanufacturing offers a promising solution, extending the lifecycle of discarded electronics and reducing environmental impact. However, warranty fraud within remanufacturing systems poses both financial and environmental threats, jeopardizing operational efficiency and undermining sustainability efforts.

This study addresses these challenges by employing ensemble machine learning models, i.e., Random Forest, Gradient Boosting Machine (GBM), and XGBoost, to enhance fraud detection in remanufacturing operations. Using a synthetic dataset designed to replicate real-world conditions, the proposed approach significantly improves the precision and accuracy of identifying fraudulent warranty claims. By capturing intricate patterns across variables such as claim amounts, service costs, and customer histories, this methodology offers a robust framework for combating fraud.

The integration of advanced analytics strengthens the integrity of closed-loop supply chains, promoting resource recovery and minimizing waste. The model's superior performance not only drives economic benefits but also aligns with broader sustainability goals, enhancing the environmental impact of remanufacturing processes.

This research highlights the critical role of machine learning in addressing the dual challenges of e-waste management and warranty fraud prevention. It underscores the high potential of technological innovation in fostering clean, sustainable practices while improving the profitability and resilience of remanufacturing systems.

A Data Envelopment Analysis Framework for Urban Ranking and Criteria Prioritization

Jin Fang¹, Yasamin Salmani², Fariborz Partovi³

¹Clark University, Worcester, MA, USA. ²Bryant University, Smithfield, RI, USA. ³Drexel University, Philadelphia, PA, USA

Abstract

This paper presents a comprehensive methodology for ranking cities and prioritizing the criteria that influence these rankings within the framework of multi-criteria decision analysis (MCDA). We develop a super-efficiency Data Envelopment Analysis (DEA) model to generate a fully ranked list of cities, ensuring robustness through data-driven and non-arbitrary weight assignments. Additionally, we employ fixed-effect and random-effects models to identify the criteria significantly impacting city rankings over time. The efficacy of these models is compared using the Hausman, Dickey-Fuller, and Breusch-Pagan tests. By analyzing large cities in the United States over 20 years, at five-year intervals from 2000 to 2020, we determine city rankings and assess the influence of various criteria. The application of fixed-effect and random-effects models allows us to evaluate the impact of each quality-of-life criterion on city rankings. Our findings reveal that cost of living, education, and income have the most substantial impacts on city rankings. These insights are invaluable for policymakers aiming to enhance urban socio-economic and environmental attributes. Furthermore, they provide businesses with critical information for strategic planning, market analysis, human resource management, and sustainable growth.

Examining Housing Instability and the Homelessness Issue in the United States: A Data-Driven Analysis

Shilpa Balan, Latika Kapoor, Kajal Bhandare

California State University, Los Angeles, Los Angeles, CA, USA

Abstract

Homelessness is a widespread issue in the U.S. with metropolitan cities facing a particularly acute crisis due to its high cost of living and housing shortages. Over the past few decades, affordable housing has gradually declined for most low-income renters in the United States. According to the U.S. Department of Housing and Urban Development, the number of homelessness has reached its highest in 2023. It is generally viewed that those who spend more than 30% of the household income toward housing costs are cost burdened. Affordable housing challenges may impact a household's budget causing hardship for necessities such as food and utilities. Housing instability is a significant barrier forming socio-economic challenges that needs to be addressed. In this research, we examine the issue of housing crisis in the United States. Our initial findings show that individuals aged 35-44 are the most affected group. Furthermore, gender-specific data show that males are more impacted with the highest homelessness rates. Ethnic disparities are equally pronounced, with African American, and Hispanic groups disproportionately impacted across the U.S.

International Investments in Artificial Intelligence: Determinants and Effects

Kraiwinee Bunyaratavej

Delaware State University, Dover, Delaware, USA

Abstract

Artificial Intelligent (AI) has received much interest in recent years due to the impact it has on both personal and business levels. Moreover, the research in AI has drawn attention to all fields of studies. In this paper, we explored the investment and the human resource management fields. We first empirically investigated the determinants of the AI investments internationally. Later, we investigated how these investments have an impact on the AI worker concentration in the countries. We used the hierarchical model for the analysis. We found that countries which have a higher level of innovation seems to attract more AI investment to the countries. The supply chain pressure variable is marginally significant. We can interpret as the supply chain pressure increases, the AI investments in a country increase. Lastly, we found that the more the AI investment in a country, we can see more AI concentration in the labor market. We hope that the results will help us better understand societal impact on AI investments.

A Review of Information Systems and Cryptocurrency Research

Subhasish Dasgupta¹, Michael Chuang², Yuan Xue³

¹George Washington University, Washington, DC, USA. ²University of Illinois, Springfield, Springfield, IL, USA. ³Elizabeth City State University, Elizabeth City, NC, USA

Abstract

Since its launch in 2009, cryptocurrencies have made a splash as a mainstream financial instrument (Duggan 2022). Financial market operations are shaped by participants' social cognitions and behaviors (SCB). The same extends to the world of cryptocurrencies. Cryptocurrency is a digital currency that serves as a medium of exchange in computer networks. Social cognition and behaviors, including customer sentiments, are evident on several social media platforms.

Several information systems studies have examined various aspects of cryptocurrency in recent years. This has included customer sentiment analysis in social media platforms and its effect on cryptocurrency prices. These technologies use cryptocurrencies as alternative payment methods and security and privacy issues in cryptocurrency, to name a few (Phillips and Gorse 2017, Derakhshan and Beigy 2019).

Our study focuses on the existing research on cryptocurrency in the information systems literature. We plan to review this literature using text-analytic methods. Our objective is to identify potential trends and identify areas of future research.

Blockchain Adoption in Supply Chain: A Trust Building Mechanism

Md Golam Kibria

Kean University, Union, NJ, USA

Abstract

Information sharing is a critical driver of effective supply chain management (SCM). Ensuring secure and transparent information sharing among supply chain partners is critical for fostering successful collaboration and strengthening relationships. Blockchain technology, recognized as a disruptive innovation, offers promising solutions to complex challenges in SCM, such as transparency, security, and trust. This study empirically examines how blockchain adoption facilitates trust-building among supply chain partners through secure and transparent information sharing. A research model, grounded in the emerging literature on blockchain and SCM, was developed and tested using survey data collected in the United States. The findings reveal that blockchain adoption enhances secure and transparent information sharing, which, in turn, fosters trust among supply chain partners. The study concludes with significant managerial and theoretical implications, highlighting blockchain's transformative potential in supply chain contexts.

Traffic Congestion as an antecedent of Data Breaches.

Anjee Gorkhali¹, Asim Shrestha²

¹Susquehanna University, Selinsgrove, Pennsylvania, USA. ²Denver University, Denver, Colorado, USA

Abstract

Data breaches generate a negative sentiment amongst customers as they are concerned about their private information, including financial and personal details (Blumenthal and McGraw, 2015), thus losing significant investment and returns. The U.S. Department of Health and Human Services (DHHS) claimed that 33 million individuals were affected in separate data breach cases in 2021 alone[\[1\]](#). Several studies identify the causes of data breaches, some suggesting technological factors (Angst et al. 2017). In contrast, others claim that a lack of information governance is the major contributor to increasing data breaches (Huang et al., 2014; Wu and LaRue, 2017). However, it is unclear whether the firms perceive traffic congestion as a risk factor. Although traffic congestion seems to be a factor outside of an organization structure, (Belland and Brant, 2018; Hennessy, 2008) indicate that it causes employees stress and is more likely to decrease employee performance. Since employee fatigue results in human error, it is more likely to increase data breaches from human errors. Human errors are a major contributor to the exponential rise in data breaches (Alvarado and Triantis, 2024; Amoresano and Yankson, 2023). Thus, we posit that traffic congestion in the hospital area is more likely to increase employee fatigue and result in increased data breaches. Using the Texas A&M Transportation Institute's Urban Mobility Reports on the traffic congestion for 494 urban areas in the U.S. and the data breach report from the Health and Human Services website, we find that traffic congestion levels in the hospital area are positively associated with data breaches.

Navigating AI Governance and Ethics: Strategies for the Future

Haleh Karimi, Alisha Harper

Bellarmino University, Louisville, KY, USA

Abstract

The rapid advancement of artificial intelligence (AI) technologies has created a pressing need for robust governance frameworks and ethical guidelines to ensure their responsible deployment. AI governance encompasses the policies and regulations that oversee the development and utilization of AI systems, while ethical considerations prioritize fairness, accountability, transparency, and the protection of human rights. Key challenges in this domain include addressing biases present in algorithms, safeguarding data privacy, and preventing the misuse of AI technologies. A practical governance framework necessitates the involvement of governments, private sector entities, academia, and civil society, ensuring the integration of diverse perspectives and promoting fairness. Critical aspects of AI governance and ethics will be explored, strategies will be identified, and the importance of collaboration and the ongoing adaptation of regulatory frameworks to keep pace with technological advancements will be highlighted.

A Conceptual Framework for Integrating Emerging Technologies into Occupational Health and Safety Management Systems

Aida Haghighi¹, Afrooz Moatari-Kazerouni², Gyeong Jun Lee¹

¹Toronto Metropolitan University, Toronto, ON, Canada. ²Widener University, Philadelphia, PA, USA

Abstract

Emerging Technologies, such as Blockchain, Internet-of-Things (IoT), Virtual Reality (VR), and Big Data play a crucial role in enhancing the implementation of Occupational Health and Safety Management Systems (OHSMS), such as ISO 45001. Such integration ensures the effectiveness of OHSMS from different perspectives. Blockchain provides secure and transparent record-keeping of accidents and compliance data. IoT platforms monitor workplaces to detect hazards in real-time. Big Data analyzes safety-related data, while VR improves safety training and emergency preparedness. This study develops a conceptual framework for integrating these technologies into OHSMS, enabling organizations to identify technological adaptation areas to enhance worker health and safety.

Developing SMARTER Robots Engaged in Robotic Interventions for Students with Learning Disabilities

Anshu Arora, Amit Arora

University of the District of Columbia, Washington, DC, USA

Abstract

Social Motivation Approach for Rehabilitation Through Educational Robotics (SMARTER**)** research project focuses on creating interdisciplinary knowledge through social-educational robotics, and targets cognitive rehabilitation in students diagnosed with learning/cognitive disabilities (e.g., Autism Spectrum Disorder or ASD) in the District of Columbia Public Schools in Washington, DC. The research objectives of SMARTER research are to: (1) examine social-behavioral relationships and human personality traits in ASD students using social motivation theory of autism; (2) investigate robotic anthropomorphism and robotic intentionality exhibited during human-robot interaction (HRI) with ASD students; (3) develop curriculum-related interactive scenarios designed for improving cognitive rehabilitation through robotic interventions targeting ASD students. The central research hypothesis of this SMARTER research is that educational-social robots may help to overcome sensory and motivational barriers encountered by the ASD individuals when they interact with humans by utilizing social motivation theory and approach in social-behavioral human-robot interaction; and educational robotics will enhance learning for K-12 students in Washington DC. The research utilizes a mix of qualitative and quantitative robotic interventions, and social-behavioral sciences methods, such as responses to surveys / questionnaires, structured content analysis, and multigroup analysis using structural equation modeling. The intellectual innovation of SMARTER project includes the creation of interactive curriculum-based scenarios based on school and university curricula utilizing ethical robotic interventions for students with learning disorders (e.g., ASD). The most significant impact of the SMARTER research is strengthening the research capacity in the interdisciplinary education, consumer behavior and robotics fields, and providing research training and opportunities for university students through educational robotics.

Keywords: SMARTER, Autism Spectrum Disorder (ASD), Social Motivation Approach, Educational-Social Robotics, Interdisciplinary Education, Consumer Behavior

Acknowledgement: The research has been funded and supported by the National Science Foundation Grants #2100934, 2406112, and 2406166.

An Ocean 97 and Green Reef Foundation Aquaculture Digital Twin Adoption Experience for HBCU's in Aquaculture with a Net Zero Waste Activity Focus

Benjamin Branch¹, Jarvis Green^{2,3}, Chris Tate⁴, Patrick Menash⁵, Kimberly Milligan⁶, Gulnihal Ozbay⁶, Micheal Casson⁶, Neset Hikmet⁷

¹Green Reef Foundation, Doha, Qatar, Qatar. ²Green Reef Foundation, Baton Rouge, LA, USA. ³Oceans 97, Baton Rouge, LA, USA. ⁴Red Hat, Boston, Massachusetts, USA. ⁵Southern University, Baton Rouge, LA, USA. ⁶Delaware State University, Dover, DE, USA. ⁷University of South Carolina, Columbia, SC, USA

Abstract

This poster will showcase the Partnership between Oceans 97, the Green Reef Foundation (GRF) and its partners to move innovation in Aquaculture. Here, www.compute.org will lead a demonstration of next generation aquaculture infrastructure that could help small business and business intelligence in aquaculture to achieve their goals. These specific outcomes target net zero waste outcomes for seafood waste lead by Oceans 97 and GRF founder Mr. Jarvis Green and his team. Such applications may then be considered and customized for use by Southern University and Delaware State University for research verification. This work discusses some zero trust aspects of implementation in this work using Red Hat solutions.

On display will be a cybersecurity implementation of Keycloak Fine Grain Authorization Permissions features to grant participants access to an OpenShift cluster, and workbenches where they can build a Smart Aquaculture project from scratch using my compute.org AI code generation technology. An OpenShift cluster was used to start testing this out as an [OpenShift Dedicated](#) cluster in Google or bare metal data center.

Digital Twinning Capital Costs (CAPx) and Operating Costs (Opx) within Process:
An Emulation in a Manufacturing Company

Douglas Hales

The University of Rhode Island, Kingston, Rhode Island, USA

Abstract

Digital Twinning is still new to the academy and industry with the largest players being IBM and Siemens. The purpose of this research is to demonstrate the difference between Simulation and Digital Twinning Opx costs in a real production process. The purpose is to show that traditional cost reporting on monthly or quarterly cycles is insufficient to inform managers about decisions facing an increasingly complex economy. The outputs of process DTs with financial data allow for much faster and efficient decision-making.

Comparative Case Study on Implementing Generative AI in Medical Practices to Ease Documentative Overburden: A Sociotechnical Systems Perspective

Sri Ramesh Eevani, Rajiv Nag

Drexel University, Philadelphia, PA, USA

Abstract

This is a comparative case study of a live implementation of Generative AI solution in 5 medical practices. We shed new light on the impact of Generative AI on various aspects such as social structures, roles, organizational processes, and technical systems of medical practices. It is well known now that increasing documentation burden on physicians has led to medical errors, patient safety concerns, and physician burnout. This study investigates the adoption and implementation of a Generative AI based clinical documentation technology in medical practices over a span of 5 months. Our data consisted of interviews, participant observations, process documentation and mapping, tracking social interactions, and analyzing textual user feedback data. The results reveal a process framework that can be generalized across medical practices, categorizing changes into social, technical, organizational, and goals & outcomes. The implementation of Generative AI has led to both tangible and intangible benefits, including the creation of a new role of Scribe to provide human oversight of AI-generated clinical documentation. Resistance and apprehensions from practice staff have impacted implementation speed and decision-making. The study emphasizes the importance of considering social and organizational process changes in the adoption of new technologies and identifies role re-forming and triadic co-creation as key concepts. The study also includes an entrepreneur's and emerging technology product implementation team's experiences of the co-creation with the medical practices. Overall, this research provides a processual framework to capture the nuances of the adoption and co-evolution of an emergent and uncertain technology.

BUILDING GENERATIVE ARTIFICIAL INTELLIGENCE (GAI)
UTILIZATION EXPERTISE AND TRUST IN GAI OUTPUT: A PARSIMONY
QUANTITATIVE MODEL APPROACH TO PREDICTORS IN THE
SOFTWARE DEVELOPMENT INDUSTRY DYNAMICS

Triumf Qosej, David Gefen

Drexel University, Philadelphia, PA, USA

Abstract

Based on an extant industry-setting business problem, we analyze the expertise-building in GAI utilization and trust in GAI output, at the agent-level scenario, in the software development industry. We intend to derive new findings and conclusions on what theory-based and industry factors contextualize and affect this process, with emphasis on the said industry. Hence, this study looks at how likely experienced software development practitioners (i.e., the agent – early adopters) are to build expertise and adopt GAI as an assistant to generate programming code (i.e., software end-product). The study adopts a between-subjects design with two software end-product complexity treatments. In the low-complexity treatment subjects were asked about using GAI to produce software that generates an automatic business email reminder. In the more complex treatment subjects were asked about using GAI to develop a universal plug-in-and-play Application Programming Interface (API) that integrates with a blockchain technology application for inter-banking payment settlements. Based on 808 respondents, the analyses find that the better model fits expertise building in GAI utilization, this latter with a strong linear relationship with adoption of GAI utilization for software development and coding. Although the interaction term familiarity*distrust requires additional study, it indicates that increased distrust reduces the positive impact of familiarity on expertise building in GAI utilization. Between-subjects high vs low complexity treatments revealed additional new findings of treatment effects, by demonstrating distinct distrust and perceived ease of use of GAI in the context of each treatment itself. There is a variance in the trust as mediating and dependent effect for the respective treatments of low vs high complexity tasks. Additional important implications, as well as study limitations and future research recommendations are also discussed.

Keywords:

Expertise-building for GAI utilization, GAI adoption for code generation, GAI for code building, AI for software engineering and development, GAI output, familiarity, trust, distrust, PEOU, PBC, intended use, complexity.

Using Natural Language Processing for Data-driven Risk Management
Haleh Valian

Biogen, Cambridge, MA, USA. Harvard University, Cambridge, MA, USA

Abstract

Quality Risk Management (QRM) in the pharmaceutical industry faces challenges in extracting pertinent information from significant amounts of unstructured text data thus hindering effective identification of potential quality risks for review and escalation. This paper proposes a new framework for Natural Language Processing (NLP)-assisted QRM to assess whether new quality issues are a recurring problem by using the unstructured issue description. In our proposed framework, we first use deep neural networks (DNNs) to extract features from unstructured pharmaceutical quality issue data. Next, we propose new data augmentation strategies and score aggregation statistics to transform the output of DNNs into a decision about whether an issue is recurring. Different DNN x data augmentation x score aggregation schemes are then evaluated empirically using 6 years of quality issue data. We found that the combination with the best overall performance is the black-box algorithm (Google's USE) x LOSO (Leave-One-Sentence-Out) data augmentation strategy x the Max-Min score aggregation scheme. Based on the performance comparison with the traditional method of identifying recurring issues using people assigned issue labels at the point of data entry, the proposed framework shows promise in offering an effective approach to improving QRM practices.

Interdependence Between Stock Return, Social Media Sentiments, and Geopolitical Risks

Xiaorui Zhu¹, Michael Lash², Kyle Li³

¹Towson University, Towson, MD, USA. ²University of Kansas, Kansas, KS, USA. ³Thomas Wootton High School, Rockville, MD, USA

Abstract

This study highlights the dynamic interplay between stock returns, social media sentiments, and geopolitical risks through a comprehensive analysis using a Vector Autoregression (VAR) model. By incorporating market variables, investor sentiment metrics, and geopolitical risk indices (GPR), the research uncovers significant bidirectional relationships among these elements. The findings reveal that positive sentiment strongly influences market returns, while negative sentiment exerts a less pronounced but significant adverse effect. Conversely, geopolitical risks negatively affect positive sentiment but positively drive investor attention to post more on the X.com. Moreover, the impulse response function (IRF) analysis reveals that S&P500 return shocks significantly increase positive sentiment in the short term, while simultaneously dampening negative sentiment, though these effects diminish over time. Geopolitical risk (GPR) shocks significantly decrease positive sentiment in the short term, with the negative effect gradually diminishing over time. This study addresses a critical gap in the literature by incorporating these three dimensions into a unified framework to understand complex financial systems and provide valuable insights for investors, policymakers, and researchers aiming to understand and manage market dynamics.

Generating Sparse and Plausible Counterfactual Explanations with Spatial Diversity and High Fidelity for Artificial Intelligence Decision Systems

Nolan Talaei, Asil Oztekin, Hongwei Zhu, Luvai Motiwalla

University of Massachusetts Lowell, Lowell, MA, USA

Abstract

Counterfactual explanations are the cornerstone of Explainable Artificial Intelligence (XAI), offering actionable insights by identifying minimal changes required to alter a machine learning model’s prediction. This paper introduces Synthesized Perturbed Anchors for Counterfactual Explanation (SPACE), a novel algorithm designed to generate diverse, plausible (aligned with data distributions), and sparse (modifying minimal features) counterfactuals efficiently. SPACE addresses key limitations of existing methods, such as lack of diversity, fidelity, sparsity, and plausibility, by synthesizing anchors in a PCA-reduced space, perturbing them directionally, and transforming them back into the original feature space. These anchors enable the exploration of distinct regions of the decision boundary, ensuring the generation of counterfactuals that provide multiple actionable options for users. Evaluated on the FICO HELOC dataset, SPACE consistently produces counterfactuals with high plausibility, sparsity, and fidelity. The implications of SPACE extend across domains and are significant, where transparency, fairness, and user empowerment are paramount. SPACE enhances trust and accountability in AI-driven decision systems by providing diverse and feasible pathways to desirable outcomes. This paper demonstrates how SPACE bridges the gap between technical performance and real-world utility, establishing it as a robust and scalable solution for counterfactual explanations.

Dynamic Vehicle Routing with Social Media Integration in Disaster Relief

Emre Kirac

Christopher Newport University, Newport News, VA, USA

Abstract

This study focuses on integrating uncertain social media data into real-time disaster relief tour planning. The proposed framework combines verified traditional data and unverified social data, updated dynamically based on verification processes. A dynamic vehicle routing model is proposed to maximize the delivery of accurate demand within time and resource constraints. The study explores various decision policies reflecting diverse emergency manager preferences, from relying solely on verified data to acting on social data with or without verification. The findings demonstrate the value of incorporating social data in specific scenarios, balancing uncertainty against the need for timely response.

Strategic Interactions in Climate Action: A Game-Theoretic Analysis of City Managers and World Leaders

Sampson Addo, Bryan Higgs

University of the District of Columbia, Washington DC, District of Columbia,
USA

Abstract

Climate change mitigation requires coordinated efforts from diverse stakeholders, including city managers and world leaders. This study utilizes game theory to analyze the strategic interactions between these actors in climate action, addressing a critical gap in understanding the interplay between local and global actors. The study seeks to answer the question: How can game theory inform climate change policy by analyzing the strategic interactions between city managers and world leaders? We aim to: (a) Develop a game-theoretic model capturing strategic interactions in climate change mitigation. (b) Identify optimal strategies and outcomes, including Nash equilibria and Pareto-efficient solutions. (c) Provide actionable recommendations for enhancing cooperation among stakeholders. Using a multi-player game-theoretic model, we identify Nash equilibria and Pareto-efficient outcomes to understand the dynamics of collaboration and competition in climate policy. Our findings indicate that stakeholders are inclined to adopt suboptimal strategies in the absence of collaboration, culminating in a “tragedy of the commons” scenario. Conversely, coordinated actions can yield Pareto-efficient outcomes that benefit all participants. Our study offers actionable insights for policymakers to enhance collaboration, promote equitable burden-sharing, and devise effective climate action strategies. Specifically, it underscores the importance of integrated local-global approaches to address this pressing global challenge. In fostering collaboration between city managers and world leaders, it is possible to achieve more sustainable and equitable climate policies. By integrating game-theoretic approaches with climate policy, we contribute to the growing body of research that seeks to understand and improve global cooperation in addressing climate change.

Enhancing Heterogeneous Treatment Effect Estimation in Personalized Marketing: Integrating Feature Embeddings with Meta-Learners and Causal Forests

Rashedul Hasan, Asil Oztekin

University of Massachusetts Lowell, Lowell, MA, USA

Abstract

Estimating heterogeneous treatment effects (HTE) is crucial for understanding variations in individual responses to interventions, particularly in personalized marketing, where customer behaviors differ significantly across subpopulations. Traditional approaches, such as the Average Treatment Effect (ATE), often fail to capture these variations, leading to suboptimal decisionmaking. Meta-learners, including the S-Learner, T-Learner, X-Learner, and R-Learner, as well as Causal Forests, have emerged as powerful tools for estimating HTE by combining causal inference with machine learning techniques. However, improving their predictive accuracy, especially in datasets typical of marketing analytics, remains a challenge.

In this paper, we propose an enhanced framework for HTE estimation by integrating deep learning-based feature embedding techniques. Feature embeddings transform categorical variables into dense, continuous representations, allowing the model to capture intricate patterns and non-linear interactions within the data. By embedding demographic, behavioral, and geographic features, we aim to improve the precision of treatment effect estimates, offering a more granular understanding of customer responses to marketing campaigns.

We evaluate our approach using real-world marketing datasets and compare its performance to traditional methods. The results show that integrating feature embeddings significantly enhances the predictive accuracy of HTE models, providing more actionable insights for personalized marketing strategies. Our work contributes to the growing body of literature combining causal inference with advanced machine learning techniques to optimize marketing interventions.

A Game-Theoretic Analysis of the Electric Vehicle Transition: Consumer, Business, and Grid Perspectives

Bryan Higgs

University of the District of Columbia, Washington, DC, USA

Abstract

The transition to electric vehicles (Evs) involves a complex interplay of decisions by consumers, businesses, and electricity providers. This paper employs game theory to analyze these interdependent decisions, focusing on cost as the primary decision variable.

- **Consumer Perspective:** Consumers face a trade-off between the higher initial cost of Evs and potential long-term savings on fuel and maintenance. This analysis incorporates factors such as driving habits, charging infrastructure availability, and government incentives in determining the optimal vehicle choice for individual consumers.
- **Business Perspective:** Businesses must decide on the type of vehicle fleet to operate (gasoline or electric) and the associated infrastructure investments (gas stations vs. EV charging stations). The analysis considers factors such as initial investment costs, operating and maintenance expenses, potential for ancillary revenue streams (e.g., retail space at charging stations), and the impact of government incentives.
- **Electricity Provider Perspective:** Electricity providers must plan for the increased electricity demand from Evs, considering the potential impact on grid stability and the need for grid upgrades. The analysis explores different strategies for grid management, including the integration of new renewable energy and the phase out of existing fossil-fueled energy.

This paper develops a multi-agent game theory model to capture the strategic interactions among consumers, businesses, and electricity providers. By analyzing the equilibrium outcomes of this game, we can identify potential market failures, explore the impact of government policies (e.g., subsidies, mandates) on the transition to Evs, and provide insights for optimizing the deployment of EV charging infrastructure and ensuring a smooth and cost-effective transition to a more sustainable transportation system.

Harnessing AI: Strategies for Enhancing Education and Research Success

Kathleen Ferris

Bridgewater State University, Bridgewater, MA, USA

Abstract

This workshop explores strategies for success in education and research during the age of AI. It provides practical insights into harnessing AI tools to enhance teaching, streamline administrative tasks, and drive innovative research. Attendees will learn how to integrate AI into course design, grading, and student engagement, as well as discover AI's potential for data analysis and research productivity. Join us to explore AI's transformative role in higher education.

Experiential Learning- Interdisciplinary Project Collaboration to Engage Students in Cost Accounting Class

Bea Chiang, Diane Bates, Kim Pearson, Monisha Pulimoon

The College of New Jersey, Ewing, NJ, USA

Abstract

As the world becomes more interconnected and complex, the role of accountants has expanded to include responsibilities in financial decisions, strategic planning, regulation compliance, and ethics. In response to the demand for reform in accounting education, educators have explored innovative methods to engage students in more effective learning experiences. One such approach is integrating interdisciplinary learning projects to develop accounting students' problem-solving skills with broader perspectives. Studies have shown a positive relationship between cross-disciplinary integration and student engagement, particularly between STEM and social science students. The interdisciplinary project studied follows the Collaboration Across Boundaries (CAB) model, which includes three components: (a) coordinated but separate courses with different instructors and goals, (b) collaboration with a community partner to address social needs, and (c) assessment of the impact on scientific literacy, disciplinary learning, and community engagement.

The research question investigates whether collaboration between computer science and accounting students improves learning outcomes in scientific literacy, accounting knowledge, and community engagement. Results show no significant differences in pre-test scores between control and treatment groups, suggesting that students in both groups rated their academic knowledge and social engagement similarly. However, both groups positively evaluated their skills in accounting, even though the differences between the two models were not fully captured by the current measures.

While the limited differences between collaborative and single-discipline groups are disappointing, there are encouraging signs that collaboration skills improved. This aligns with anecdotal evidence from students and faculty. We suggest that learning to collaborate across disciplines is an essential academic skill and recommend future research to develop better measures for evaluating collaboration.

What should we be teaching in UG Introduction to Information Systems?

Neil Desnoyers

Saint Joseph's University, Philadelphia, PA, USA

Abstract

The last several semesters of teaching F2F undergraduate Introduction to Information Systems have not been a satisfying experience for this instructor. The current version of the course is structured as something of a “history of technology” course. While rather interesting to this instructor, this version of the course has not resonated with the students. An example of student feedback along this line is the following student evaluation comment from the Fall 2024 semester: “...this course I feel is very outdated...”. Our college has decided that this course, beginning in Fall 2025, will be the primary course in which undergraduates are introduced to the promises and pitfalls of Artificial Intelligence (AI). What should that look like? Possibilities will be discussed.

Enhancing Decision Support Skills in Engineering Graduate Students through Game Theory

Bryan Higgs

University of the District of Columbia, Washington, DC, USA

Abstract

This paper investigates the value of integrating a game theory course into an engineering graduate program to bolster student decision-making abilities specifically within the context of climate change. The inherent uncertainties and complexities of climate change, coupled with the involvement of multiple stakeholders and competitive pressures, demand a nuanced approach to engineering solutions. Traditional engineering curricula may not adequately equip students with the tools to navigate these multifaceted challenges effectively.

Game theory offers a robust framework for analyzing strategic interactions, modeling both competitive and cooperative behaviors, and guiding optimal decision-making under conditions of uncertainty and competing interests. By introducing core game theory concepts, such as Nash Equilibrium, dominant strategies, and cooperative game solutions, students can:

- Cultivate strategic thinking and decision-making prowess: Analyze situations from diverse perspectives, anticipate the actions of other stakeholders (governments, industries, communities), and make informed choices that carefully consider potential consequences.
- Enhance problem-solving capabilities: Apply game theory principles to model and resolve intricate engineering problems related to resource allocation, network design, control systems, and market competition, all within the context of climate change mitigation and adaptation.
- Sharpen critical thinking and analytical skills: Develop the ability to abstract complex climate change challenges, identify key players and their respective incentives, and formulate effective and strategic solutions.
- Gain valuable insights into real-world engineering challenges posed by climate change: Explore the application of game theory to address contemporary issues such as the integration of renewable energy systems, the development of resilient infrastructure, and the mitigation of climate-related cybersecurity threats.

This paper will delve into the specific learning objectives of a recently implemented game theory course for engineering graduate students, outline the pedagogical approaches employed, and discuss the observed impact on student learning outcomes. Furthermore, the paper will address the challenges of integrating game theory into the engineering curriculum.

What Do Students Say About HBCUs? A Preliminary Text-Mining Analysis

Cretson Dalmadge, Shih Chou

Winston-Salem State University, Winston-Salem, NC, USA

Abstract

Over 180 ago, historically black colleges and universities (HBCUs) were founded to provide access to higher education to African American students (Kennedy, 2012) and a means to address racial and other inequities in society through education (Crewe, 2017). Although HBCUs have their unique mission and role in the landscape of higher education, they often face more challenges, such as decreased demand and retention.

Without a doubt, the unique role that HBCUs play in higher education has attracted a considerable amount of scholarly attention, which can be broadly classified into understanding the strengths and challenges of HBCUs and exploring student outcomes at HBCUs. Even with the existing scholarly progress, comparatively less is known about how students at HBCUs perceive the campus and institution from various domains.

To address the aforementioned gap in the literature, this study attempts to understand students' ratings and textual reviews of HBCUs using text mining to extract key themes from students' textual reviews left on RateMyProfessors.com over a 10-year period ranging from January 2014 to December 2023. A total of 2,911 individual responses from 94 HBCUs in the U.S. were collected and analyzed using the K-means clustering algorithm, which resulted in 5 clusters. In particular, cluster 1 focuses on academic-related attributes at the institutional level (e.g., program offerings). Cluster 2 emphasizes counseling and mental health services on campus. Cluster 3 revolves around student life on campus. Cluster 4 centers on teaching, curriculum, and academic support. Cluster 5 concentrates on campus administration.

Online Asynchronous Case Teaching in an AI World

Ellen Thomas

New Jersey Institute of Technology, Newark, NJ, USA

Abstract

Using the case study method to teach MBA students has many advantages. For example, it allows students to grapple with complex real-world situations while developing critical thinking and problem-solving skills. Online asynchronous MBA courses introduce challenges in that the benefits of case discussion and collaboration are almost impossible to achieve. Set this online asynchronous MBA course in an AI world of plagiarism, and it is nearly impossible to manage. What more if the class is large?

Harvard Business School Publishing recently published a poll asking instructors, “Are you worried about students using AI in your case classes?” Fifty-nine percent responded no, while forty-one percent responded yes; they are worried. Concerns included loss of critical thinking skills and plagiarism; 65% of students admit to using AI to cheat. On the other hand, those not concerned expressed benefits centered on class interaction.

How should these concerns and benefits be addressed? Instructional design. Augment case studies with online asynchronous interactive presentations and discussions.

There are many tools and techniques available to instructors. The one presented here is VoiceThread, an interactive presentation tool (<https://voicethread.com>). VoiceThread allows students to create slideshows using images, videos, audio, drawings, and more. When uploaded to the course, it is made available to classmates for comments, feedback, and collaboration.

This presentation will discuss how this interactive tool can supplement online asynchronous case study teaching to ensure student engagement and understanding. Assignments designed for an MBA Marketing Strategy course are reviewed, and findings related to authenticity and learning in a mid-sized public East Coast university will be shared.

Enhancing Operations Management Education with AI: Engaging Students with Innovative In-Class Activities

Jonathan Jackson

Providence College, Providence, RI, USA

Abstract

Incorporating artificial intelligence (AI) into classroom activities offers innovative ways to engage students and deepen their understanding of key concepts. This presentation outlines my approach to integrating AI into an undergraduate Operations Management course through chapter-specific activities. Students explored use cases across industries, brainstormed solutions to complex problems, and constructed debate arguments on controversial topics.

These activities enhanced engagement, fostered critical thinking, and provided hands-on experience with AI tools, preparing students for real-world applications. Feedback showed students found the activities engaging and beneficial for solidifying their learning.

This presentation will share activity designs, outcomes, and practical tips for educators interested in leveraging AI to enrich their teaching, regardless of discipline.

Innovations in the Undergraduate Spreadsheet Modeling Course

Ross Malaga, Rreze Zejnullahi

Montclair State University, Montclair, NJ, USA

Abstract

Introduction to Spreadsheet Modeling is a required class for all business students. The course struggled with high student failure rates. Over a period of 2 years the course was redesigned. The modality shifted from asynchronous online only to asynchronous online with required lab attendance. Success coaches monitored student performance and intervened when needed. The course was tied to the MS Excel certification exam and a course fee was put in place so each student can take the exam twice. Failure rates dramatically decreased, while performance on the certification exam exceeded expectations. The change process and lessons learned will be discussed.

Learning from Experience: A ChapGPT application in a classroom

Anil Aggarwal, Cong zhang, Danielle Fowler

University of Baltimore, Baltimore, MD, USA

Abstract

Artificial Intelligence (AI) is revolutionizing the world... Data is being collected at a tremendous rate which needs to be filtered, organized, analyzed and implemented for strategic advantage. Many systems are emerging that are learning and just doing that. It creates opportunities as well as challenges. Businesses are demanding students who are competent in searching, managing and analyzing data. Universities are meeting these challenges by providing “data” focus courses. Though AI is huge and consists of many different areas like robotics, medical diagnosis, robotic surgery etc. but our focus here is on appropriate “data” collection and filtration. As part of continuous improvement, information literacy was included as part of the first introductory information course at the undergraduate level. This experiment is continuing and has also happened over several semesters. Instead of re-inventing the wheel, academicians can use our experience as the starting guide.

Information literacy was used as part of an assignment. The assignment requires students to search three sources: ChatGPT or any such AI-based tool (some students used BARD). Any library and the Internet. Students had to select articles from each of these sources and evaluate them in terms of several factors like currency, bias, grammar, content, etc. Students were required to use keywords, using any algebraic combinations. Though students were not required to compare the tools but some provided their assessment of difficulties in each of the three tools.

Three professors were teaching several sections and collected comments and discussed about continuous improvement using some of the following strategies: (this is not a complete list)

- Increase the focus from micro to macro by focusing on (LLM Large language model)
- Discuss GPT (generative pre-trained model)
- Pick one (i.e., ChatGPT) and discuss its working
- Understand its protocol to get optimized output
- Discuss suitability factors
- Assess article’s suitability (filtration)

Exploring A Constructive Use of ChatGPT in the Classroom

Yang Li, Ram Misra

Montclair State University, Montclair, NJ, USA

Abstract

This study explores the adoption of ChatGPT in improving teaching effectiveness. Students from multiple sessions of graduate class (MBA) on Supply Chain and Operations Management decided are required to use ChatGPT to generate a term paper outline and develop it into a full paper. The preliminary results indicate that students found ChatGPT helpful in generating their own papers. Meanwhile, our analysis shows beginners of ChatGPT have limited capacity to generate high-quality content based on ChatGPT generated outline. In addition, text mining is conducted to compare the readability and information density of ChatGPT-generated content and student-generated content.

An Analysis of the Impact of Student Perceptions of an ERPsim Game on Their Performance

Sung Shim

Seton Hall University, South Orange, New Jersey, USA

Abstract

ERPsim is a set of computer-based simulation games designed to teach students about SAP enterprise resource planning (ERP) systems. In these games, participants manage a virtual company within a competitive market environment, using the ERP system to make business decisions. As ERPsim games become increasingly popular in educational settings for teaching ERP systems and related business processes, there remains limited understanding of how students perceive these games, particularly in terms of their perceived usefulness and ease of use, and how these perceptions influence student performance. Research in the broader field of information systems suggests a strong link between user perceptions and performance outcomes, but few studies have explored this relationship within the context of ERPsim games. This study seeks to fill this gap by examining how students perceive the usefulness and ease of use of ERPsim games, with a focus on key system features, and how these perceptions correlate with student performance during gameplay. Data was collected through a survey of college students enrolled in a course on ERP systems, who used the ERPsim Distribution game. The findings of this study can provide valuable insights for educators and developers of ERPsim games, helping to enhance game design and instructional strategies to improve learning outcomes in ERP system education.

Redefining Academic Support: The Synergy Between GPT Assistants and Embedded Librarians in Online Classrooms

Ronda Mariani, Michael Coffta

Commonwealth University of Pennsylvania, Bloomsburg, PA, USA

Abstract

The rising enrollment in courses and the growing acceptance of online learning underscores the necessity for innovative support systems that enhance student engagement and academic success. Integrating Generative Pre-trained Transformer (GPT) assistants alongside embedded librarians offers transformative potential for virtual classrooms, addressing scalability, engagement, and academic support challenges. This paper explores the complementary roles of GPT and embedded librarians in redefining educational support systems.

GPT assistants excel in real-time responses and resource recommendations, automating repetitive tasks and expanding accessibility (Bastola et al., 2023; Sajja et al., 2023). In contrast, embedded librarians contribute to critical student support and ensure a human-centered approach to academic guidance (Carroll & Mallon, 2021). Through a case study, we examine the practical applications of this hybrid model and its ability to enhance student engagement and learning outcomes. Findings highlight how the integration of GPT allowed librarians to focus on personalized support, aligning with Sharma and Yadav's (2022) assertion that AI tools can complement rather than replace traditional roles in education. Student feedback underscored the benefits of this dual-support model, including increased satisfaction and improved resource accessibility.

This paper provides actionable recommendations, encourages interdisciplinary collaboration, and discusses potential ethical considerations for AI use.

Alice in Borderland: The King of Diamonds Classroom Activity

Michael Douglas

Millersville University, Millersville, PA, USA

Abstract

Alice in Borderland is a Japanese comic/manga book series that was made into a popular Netflix TV Series. The main character and friends are transported to a fantasy world where they need to complete contests to survive in a Hunger Game or Squid Game style battle royale. The contests are based on a deck of playing cards. With the higher the card value, the harder the contest. For the 3 of Clubs, a partial task is to solve a math problem. In a harder later challenge, the King of Diamonds challenge, the players are presented with a challenge based on the idea of the Keynesian Beauty Contest.

In volume 15 of Alice in Borderland, the “Game: Beauty Contest” is introduced. This presentation shows how to operationalize the Alice in Borderland Beauty Contest to the business classroom, particularly an in-person MIS course. The topic being covered in class is social media, and particularly the website Stocktwits.com. Why would professional investors look at stocktwits.com or r/wallstreetbets on Reddit? Why would people with 1 share of a company stock spend hours listening to a company’s quarterly results conference. In the 1930s John Maynard Keynes wondered why stocks move the way they did. Today, we still have the same principles at play.

This presentation will present the rules of the contest to allow educators to run this exercise in the classroom. Finally, an after-activity discussion of select behavioral finance and social media is conducted to explain recent meme activity with stocks such as GameStop and AMC using the Keynesian Beauty Contest theory. This is excellent teaching and learning activity relevant to the modern post-covid face-to-face student.

A META ANALYSIS OF 30-DAY UNPLANNED READMISSION ESTIMATION MODELS

Fatma Pakdil¹, Maling Ebrahimipour², Drew Zhang², Benjavan Upatising³, Yuehwern Yih⁴, Mahtab Kouhizadeh²

¹Eastern Connecticut State University, Willimantic, CT, USA. ²University of Rhode Island, Kingston, RI, USA. ³Center for Innovation and Transformational Change Department, Umass Memorial, Worcester, MA, USA. ⁴Purdue University, West Lafayette, IN, USA

Abstract

Hospital performance is frequently assessed using key metrics such as charges, length of stay (LOS), and readmission rates, with the 30-day unplanned readmission rate standing out as a critical indicator of care quality. Often considered an “adverse outcome,” unplanned readmissions disrupt patient expectations and continuity of care and impose significant financial and operational burdens on healthcare systems. The Centers for Medicare and Medicaid Services (CMS) estimates that approximately \$17 billion is spent annually on avoidable readmissions, which account for 20% of all hospital readmissions within 30 days of discharge. While unplanned readmissions are particularly concerning for patients with chronic diseases, research suggests that many are preventable through improved inpatient and post-discharge care. The current study systematically examines readmission prediction models and methodologies discussed in the literature to address this issue. A review of 2,598 scholarly articles published between 1991 and 2024 categorizes these methods into three primary groups: 1) multivariate statistical techniques, 2) machine learning approaches, and 3) hybrid methodologies. The study also investigates the datasets used, analyzing their scope—ranging from single-facility to nationwide datasets—and geographical distribution. Furthermore, variables related to social determinants of health and comorbidity indexes are examined to understand their influence on readmission predictions. A comparative analysis is performed to assess the managerial and clinical implications of various estimation models. This research provides a comprehensive overview of readmission prediction methods, highlighting their strengths and limitations, and offers actionable insights for improving healthcare quality and operational efficiency, empowering the decision makers to make positive changes.

Customer and Employee Perceptual Congruence in Service Co-production: A Simulation Study

Ahmet Ozkul

University of New Haven, West Haven, CT, USA

Abstract

Perceptual congruence research argues that when a customer and a front-line employee have similar perceptions of a co-produced output, service quality improves. We conduct a computer simulation study based on an analytical model describing perceptual congruence in a dyadic customer-employee relationship using a constant elasticity of substitution (CES) function with customer and employee perceived efforts as input and perceived co-production as output. When analytically examined, the CES function results in four service environments in which perceptual congruence is feasible. This study focuses on finding the factors that lead to these environments. Initial simulation runs use such experimental factors as mean and standard deviation of efforts, level of correlation between efforts, and number of periods the past efforts remembered.

The role of social networks in the perceptions of health professionals about new technology: The Example of ESOGÜ Health and Practice Research Hospital

NESET HIKMET¹, MENDERES TARCAN²

¹University of South Carolina, Columbia, South Carolina, USA. ²Eskisehir Osmangazi University, Eskisehir, NA, Turkey

Abstract

Aim: It was conducted to examine the role of social networks in the perceptions of health professionals about new technologies.

Objective: The population of this research, which was conducted with the cross-sectional research method, consists of health professionals working in the operating room of ESOGÜ Health and Practice Research Hospital. The “An Update and Streamlined Technology Readiness Index: TRI 2.0” scale was applied to determine the perceptions of the sample group of 44 people selected from this universe towards new technology. In order to create the social network structure of the sample group, three questions were used to create the positive network, the counseling network and the negative relationship network. The obtained data were analyzed with SPSS 20.0 and UNICET 6.6 program.

Methods: 43,2% of the participants were male and 56,8% were female. 34,1% consists of physicians, 54,5% nurses, 4,5% operating room technicians, 4,5% health technicians, 2,3% radiology technicians. 65,9% of the participants work between 1-10 years, 11,4% between 10-20 years, 20,4% between 20-30 years, and 2,3% of them work for 30 years or more. Average distance values are 1.514 for positive relationship network and 2.552 for client relationship network. The number of components in the social network is 43 for the positive network and 56 for the client network.

Results: It has been seen that the social network structure, in which the demographic characteristics of health professionals play a role, plays a role in the perceptions of the actors regarding technological innovations. Positive and client relationship ties within social network structures have shaped the flow of information and resources according to the internal degree of centrality, closeness and in- betweenness they receive.

Key Words: Health technologies, social networks, centrality, betweenness, fairness

Nurse Staffing in Rural Hospitals and Hospital Performance

Dinesh R. Pai¹, Esmaeil Bahalkeh²

¹Penn State Harrisburg, School of Business Administration, Middletown, PA, USA. ²University of New Hampshire, Health Management and Policy Department, Durham, NH, USA

Abstract

Nursing is a significant component in influencing the quality of care in hospitals, directly impacting patient outcomes. A growing body of research on nurse staffing and patient outcomes has demonstrated that higher levels of nurse staffing are associated with improved quality of care and operational performance. Using 2000-23 data on Pennsylvania's rural hospitals, this study investigates the impact of nurse staffing patterns on the performance of rural hospitals.

A Validation Analysis of the Resident Matching Algorithm

Kartikeya Puranam, Briance Mascarenhas

Rutgers University, Camden, NJ, USA

Abstract

The intent of the residency algorithm is to match medical graduates and programs based on their ranked preferences to improve applicant-program fit and patient care. The algorithm has been used for more than a generation without an external validation that can provide feedback to improve it. This validation study runs the algorithm 2300 times using randomized applicant and program ranking inputs. The study finds that the algorithm's matching outcomes are related to the input preferences of programs but unrelated to those of applicants.

Creating Resilience for Minorities and Female Workers, and the Role of Flexibility in Work Environments: A Mixed Method Study

Dawn Adams-Harmon, Jung Ah (Claire) Yung, Aleisha Lewis

Kean University, Union, NJ, USA

Abstract

Prior to the COVID-19 pandemic, remote and hybrid work options were scarcely available. Where applicable, workers experienced remote and hybrid flexible work arrangements, during the pandemic. Post-pandemic, many companies are pivoting and rescinding remote and hybrid options and requiring in-office attendance. The impetus of this research was to ascertain future workers' preferences for remote, hybrid, and live work situations, as current students have become accustomed to flexible work and school standards. This mixed method research was conducted with 289 undergraduate business students studying online and on-campus in an urban University in New Jersey, within the United States; in May and September of 2023. Open-ended and likert scale survey questions were given anonymously, to business students with a 72.3% response rate. The following quantitative hypotheses were researched and should significance:

H1: Gender (Women v. Men)-Female students may perceive workplace flexibility more importantly than male students.

H2: Racial/Ethnic Minority-Racial/ethnic minority students may value workplace flexibility more than white students.

H3: Intersectionality-Racial/ethnic minority female students may perceive workplace flexibility as more important than other students.

In terms of the qualitative findings, the results uncovered that female and racial/ethnic minority students regard flexible work arrangements such as hybrid work, more importantly than others. Females and racially diverse employees desire greater flexibility to maintain work-life balance. Lastly, female students perceive work-life balance very differently from male students. The benefits to workers include decreased commuting costs, increased commitment, and decreased stress; while employers benefit by keeping attrition levels low, increasing employee engagement, decreasing office space and associated overhead costs, while maintaining productivity. The implication for employers is that flexible work options should continue to be offered and maintained instead of being eliminated, especially if a corporation desires to increase the numbers of females and racially diverse employees at their companies.

CrowdStrike Outage: The Social Responsibility of Technology
Nancy Sullivan¹, Ellen Belitzky²

¹Exact Changes, River Vale, NJ, USA. ²University of New Haven, West Haven, CT, USA

Abstract

Publicly known internet outages or security weaknesses have been thought to cause financial loss through information or identity theft. On July 19, 2024, a technology tsunami caused massive damage resulting in unprecedented broader impact [9]. The CrowdStrike outage demonstrated that internet dependence could be life-threatening, global and without borders. Loss from technology induced events before the CrowdStrike outage have been massive, but relatively contained to a specific organization, government body or database. Incidents in the past have led to cybersecurity as a multi-billion-dollar industry with a substantial year over year growth rate. Public awareness of the criticality of internet operations made a step change during the CrowdStrike outage. The incident cast a spotlight on the social impact of technology. With nearly half of Fortune 100 companies reliant on CrowdStrike's cybersecurity platform [8], the failure illuminated the systemic fragility of our increasingly interconnected digital infrastructure.

This exploratory study investigates the social responsibility of technology from the perspective of behaviors among governments, technology services, corporate entities and private citizens. Understanding the CrowdStrike outage is important beyond the protection of financial assets of individual organizations and extends to broader repercussions for societal impact. The study finds that the CrowdStrike outage simultaneously caused financial loss, urgent health care disruptions, national security risks, communication network failures, and air safety vulnerabilities at massive levels of magnitude. The damage was not limited by the typical powers of containment such as location, border, industry, sovereignty, or capacity of global reach. Threats of such a nature must be considered, managed and regulated on the same scale as a global public risk across all sectors of health services, safety services and financial markets. This requires changes in behavior for usage of data, generation of data, and storage of data by governments, organizations and private individuals.

Modern Slavery and Human Trafficking in Supply Chains

Wojtek Wolfe, Srinivasa Puranam

Rutgers University, Camden, NJ, USA

Abstract

Modern slavery and human trafficking are very prevalent in supply chains. Advancements in technology have yielded new methods and opportunities for identifying and impeding human trafficking. For example, data mining techniques and machine learning algorithms can detect patterns and trends within extensive data sets, making it possible to uncover potential trafficking networks and high-risk areas. This project aims to understand how technology simultaneously facilitates and obstructs human trafficking within a geopolitical understanding.

In focusing on the intersection of geopolitics, the use of technology, and modern slavery, the first objective aims to lay a solid groundwork for the research through a detailed literature review and environmental scan. This crucial step will help pinpoint the existing knowledge, gaps, and potential opportunities in the field. Delving into the business models, supply chain intricacies, and the interplay of political, legal, and regulatory systems highlights the necessity of grasping the wider context where technology and human trafficking converge by looking at human trafficking through the lens of the supply chain – where human trafficking begins, who are the key players, and how it gets to its final consumers.

Customer Relationship Management for Intangible Goods: The Effects of Customer Commitment, Engagement, Responsiveness, and Reactance on Customer Rejection

Ameya Badve¹, Ramakrishnainh Bezawada¹, Karthik Sridhar²

¹Baruch, NYC, NY, USA. ²Baruch, nyc, ny, USA

Abstract

In a market driven by technology and information, the consumption of intangible goods has gone up. Customer relationship management for intangible goods is quite different from tangible goods. This article introduces a conceptual framework of customer relationship management for intangible goods that captures customer relationship using a novel construct called customer rejection and hypothesizes the effects of customer commitment, customer engagement, customer responsiveness, and customer reactance on customer rejection. The authors empirically test the hypotheses using data from a mobile app consumption. They find that while the risk of customer rejection is mitigated by the higher levels of customer commitment, customer engagement, and customer responsive, such risk is increased with a higher level of customer reactance. However, the increased risk of customer reactance can be managed using active customer engagement and customer responsiveness. Finally, exploring the different types of customer rejection, the authors provide strategies for hard and soft customer rejections based on product usage and customer lifetime. This research provides theoretical and managerial insights into the customer relationship management for intangible goods.

Consumer Perspective on Forced Labor: Product Price Tolerance

Estefany Galdamez

New Jersey Institute of Technology, Newark, NJ, USA

Abstract

This Research paper aims to provide a Systematic Literature Review of peer-reviewed articles on the consumer perspective on firm misconduct, focusing on forced labor. First, it gives an overview of firm misconduct from the consumer perspective and then pivots to a form of misconduct: forced labor. This literature review will identify gaps in the research and then propose a survey study that will explore the relationship between the consumer's knowledge of forced labor and the pricing tolerance of a product.

Keywords: Firm Misconduct, Consumer Perception, Forced Labor, Government Involvement, Price Tolerance, Ethical Products

Reactions to Competence- versus warmth-based brand transgressions: Role of political orientation

Feisal Murshed¹, Tilottama Chowdhury²

¹Kutztown University of Pennsylvania, Kutztown, PA, USA. ²Quinnipiac University, Hamden, CA, USA

Abstract

The current research offers a novel perspective of potential consumer reactions following a brand transgression. Specifically, we examine how consumers' political ideology may systematically influence their response to competence-based versus warmth-based brand transgressions. In two online experiments, 116 and 76 participants were exposed to a series of brand transgression real-time (vs. hypothetical) examples of brand transgressions—either competence- or warmth-related—and their brand attitudes were measured. The results reveal that, in the case of competence-based transgression, participants with a liberal orientation exhibited significantly less positive brand attitudes, while no such effects of political orientation were observed for warmth-based brand transgressions. The study concludes with a discussion of the implications and suggestions for future research.

Bridging the Gap in Marketing Operations: The Agile Campaign Managed Customer Framework.

Amirhossein Saghezchi, Seung-Lae Kim

Drexel University, Philadelphia, Pennsylvania, USA

Abstract

This research introduces an innovative approach called Campaign Managed Customer (CMC), which adapts the fundamental principles of Vendor Managed Inventory (VMI) from inventory management to the marketing domain. VMI is a collaborative supply chain strategy where the supplier assumes responsibility for maintaining the buyer's inventory levels within predefined upper and lower thresholds. By granting suppliers significant autonomy over replenishment decisions, VMI ensures the efficient availability of goods at the buyer's location and minimizes information distortion.

Similarly, many organizations lack a streamlined and effective method for determining the optimal timing for launching marketing campaigns to address declining customer numbers. Current practices are often hindered by delays in communication between organizations and marketing agencies, and incomplete or imprecise data regarding the current customer base. This study aims to develop an agile CMC framework in which marketing campaigns gain direct access to customer metrics. Under the CMC model, campaigns would be triggered automatically when customer numbers fall below a specified threshold, thereby reducing the gap between recognizing the need for a campaign and initiating it. Based on CMC assumptions, the organization would set campaign budgets, while the marketing agency would retain control over the campaign's duration, location, and type.

By adopting the CMC approach, this research seeks to bridge the gap between campaign planning and execution, enhancing responsiveness and efficiency in marketing operations. This novel approach promises to improve customer retention and streamline coordination between organizations and their marketing agencies.

Comparing virtual reality and website as shopping medium: Perspective from the S.O.R model

Eklou R. Amendah¹, Archana Kumar²

¹University of Southern Maine, Manchester, NH, USA. ²Montclair State University, Montclair, NJ, USA

Abstract

The objective of this study is to examine the influence of virtual reality and website stimuli on users' decision-making processes when shopping in these two channels. Scholars predict that virtual reality will become a prevalent shopping method (Billewar, 2022; Luna-Nevarez & McGovern, 2021); however, there are few studies that investigate consumer perception of virtual reality retail (VR) as a shopping medium. Like online shopping, virtual reality retail shopping is predicted to be widely adopted by consumers because of its potential benefits for both consumers and retailers (Luna-Nevarez & McGovern, 2021; Peukert et al., 2019). It is therefore important to investigate the factors that determine consumers' intention to consider virtual reality retail as a shopping medium and to examine whether some of the factors that determine online shopping apply to virtual reality shopping also.

Using the stimulus, organism, and response model (S.O.R), the authors first test the relationship between virtual reality stimuli and perceived risk, perceived value, and trust. Second, the relationship between virtual reality stimuli and emotion, enjoyment and flow state is tested. The same relationships are tested in the context of online shopping (the relationship between website stimuli and perceived risk, perceived value, and trust; the relationship between website stimuli and emotion, enjoyment and flow state). The authors also tested the relationship between perceived risk, perceived value, trust, emotion, enjoyment, flow state, and intention to use.

Experiment design is used to test these hypotheses. One group of participants used an apparel store as the online shopping medium while the second group shopped in a virtual reality store.

The study helps delineate the components of websites and virtual reality in the retail environment. The study sheds light on the way virtual reality as a shopping channel is designed to help infuse flexibility in retail operations and efficiency in management activities.

Unlock Success: How Organizational Resilience Bridges Resources and Performance

Huiling Liu

Arkansas State University, Jonesboro, AR, USA

Abstract

The purpose of this study is to capture the collective understanding of the role of organizational resilience in strategic resource management, emphasizing on the drivers of organizational resilience and its ability to leveraging resources for competitive advantage. Simultaneous Equation Modeling is used to estimate the moderated mediation model with a sample of 46,596 firm-year observations across 175 manufacturing industries from 1990 to 2021. Analysis results show that physical assets and unabsorbed slack, rather than absorbed slack, are drivers of organizational resilience, and that organizational resilience is positively related to abnormal performance, and such effects are strengthened by environmental dynamism.

Model Templates for Optimization Model Development in Excel

Janos Pinter

Rutgers University, Piscataway, NJ, USA

Abstract

In many courses taught at business schools as part of the management science curriculum, Excel is the advocated modeling platform of choice. Due to the specific features of working in Excel, this creates a challenge, but also an opportunity to introduce students to structured model development. In this talk, we present model development templates for a number of standard optimization model types introduced in the classroom. Structured modeling supports communicating models to the Excel Solver in a concise and transparent manner. The active use of model development patterns helps students to develop insight, and to apply the acquired modeling skills also when working in other optimization modeling environments.

Optimal Charging Control for Autonomous Electric Ride-Sharing Fleet

Ehsan Mahyari

University of Scranton, Scranton, Pennsylvania, USA

Abstract

The integration of autonomous driving technology with electric ride-sharing fleets (AERF) offers promising solutions for sustainable urban mobility. Autonomous vehicles (Avs) reduce operational costs, enhance safety, and improve fleet efficiency. When combined with ride-sharing, AERFs can alleviate congestion, lower emissions, and decrease parking demand. However, managing AERFs introduces complex challenges, particularly in optimizing vehicle deployment, ride demand, and charging schedules. Unlike traditional fleets, Avs can autonomously navigate to charging stations, but determining the optimal timing for recalling vehicles for charging—while balancing ride demand—remains a critical issue.

Current electric vehicle charging management systems (EVCMS) monitor charging processes but lack advanced decision support tools for fleet optimization. Existing research focuses on minimizing charging costs for vehicles arriving at charging stations, assuming vehicle arrivals are independent. This assumption does not apply to AERFs, where vehicle recall can be controlled by operators based on real-time factors such as ride demand and state-of-charge.

To address this gap, I propose a simulation optimization model to develop near-optimal policies for recalling Avs for charging. The model uses real-time data on vehicle charge levels and incorporates uncertainties such as energy consumption, electricity prices, and charger availability. The proposed approach aims to minimize energy costs while ensuring that ride availability is maintained. This model provides a foundation for more efficient and sustainable management of AERFs, enabling fleet operators to make data-driven decisions that balance energy efficiency, operational costs, and service quality.

A clustering-based approach for items storage and allocation problem

Reza Nafar, Satyaveer Singh Chauhan, Salim Lahmiri

Concordia University, Montreal, Canada

Abstract

A supply chain aims to deliver superior services while maximizing cost efficiency. A warehouse plays a crucial role in logistics systems and significantly contributes to achieving supply chain objectives. However, inefficient warehouse operations can quickly erode these benefits. Consequently, improving storage efficiency and reducing customer response times have become vital for enhancing warehouse performance. With limited space relative to the number of items, effective item allocation and picking become increasingly challenging.

In this study, we focus on storage and picking operations. As demand fluctuates, the location and capacity of each item must be re-evaluated. Solving the entire problem optimally is challenging, so we approach it using cluster and association analysis-based approaches. Two clustering techniques – pairwise comparison clustering, and K-means clustering, are employed to develop solutions to the picking and storage problem. The performance of both methods is compared to the random storage policy currently used by the participating company. Both clustering algorithms consider item associations, grouping items based on how frequently they appear together in customer orders. The next step applies assignment techniques to place the clustered groups in each aisle, minimizing the total number of aisle visits and ultimately reducing picking distance. By emphasizing item associations, our model is particularly suited for orders with multiple items in the modern retailing sector. It also significantly reduces picking distance compared to the random storage method. A case study and the numerical performance of the approach will be presented.

Cyclic Routing Problem with Time Separation Constraints

Gang Li¹, Anantaram Balakrishnan²

¹Bentley University, Waltham, MA, USA. ²The University of Texas at Austin, Austin, TX, USA

Abstract

We study a cyclic routing problem in which a vehicle must visit network links over multiple time periods, starting and ending at a base location in each cycle. The problem includes constraints on the start and end points, visit frequency, and time separation between consecutive visits to the same link, both within and across cycles. These constraints ensure efficient vehicle use by preventing resource waste from revisiting too soon and service disruptions from revisiting too late. This problem originates from the rail track inspection scheduling problem, which, in addition to the above constraints, involves specific engineering requirements. In this paper, we propose an innovative two-level framework that separates application-specific requirements from a high-level model designed to minimize cycle time while meeting common routing and separation constraints. This framework enhances the model's versatility, enabling it to efficiently solve not only the railway application but also a broad range of similar routing and scheduling problems. To address the challenge of large model size due to separation constraints, we develop a customized solution approach incorporating model reduction, strong inequalities, and optimization-based heuristic methods. We tested our model and solution strategies on a large railway network's geometry car scheduling project, where they significantly outperformed both the previous planning model and commercial software solutions.

The Effect of Consumer Attributions and Relationships on Return Recovery Outcomes in the Apparel Industry

Dan Liu¹, Guangzhi Shang²

¹James Madison University, Harrisonburg, Virginia, USA. ²Florida State University, Tallahassee, Florida, USA

Abstract

Despite the potential challenges associated with product returns, previous research has unveiled a significant revelation: when retailers offer superior return services, a return experience can transform into an opportunity to fortify consumer relationships and encourage repeat purchases. However, existing research has predominantly treated returns as a uniform concept, neglecting to categorize them or investigate the distinct effects of each type. This study aims to bridge this research gap by initially classifying return types based on reasons reported by customers. Subsequently, the research explores how different return types can influence customers' likelihood of making future purchases. The findings uncover a significant variation in the impact of various returns on repurchase. Notably, returns arising from sizing problems exert a positive influence on repurchase, potentially due to their self-attributed nature. Conversely, returns prompted by liking issues have a negative impact owing to their seller-attributed nature, while quality-related returns involving third-party logistics do not demonstrate a significant effect on repurchase. Moreover, the study reveals a positive moderating effect of consumer relationships on the return-repurchase dynamic. This paper provides insights from the consumers' perspective, addressing the scarcity of customer viewpoint in the return service recovery literature. It delivers valuable and distinctive insights for businesses seeking to optimize their consumer return operations, refine strategies for resource allocation, and enhance their customer relationship management practices.

Analyzing Bikeshare User Preferences Across Urban Zones: A Case Study of San Francisco's Zoning Influence on Ridership

Kaverinathan Dhanapal, Sanuj Kumar

New Mexico State University, Las Cruces, New Mexico, USA

Abstract

The relationship between bike-share user behavior and city zoning conventions as an area of inquiry is relevant for bike-share operators from an operational effectiveness perspective. Recently, some have highlighted the collaborative design of public transport and bike-share systems based on city zoning conventions. This study examines the relationship between bike-share users' behavior and city zoning conventions. Also, it helps to identify the underlying links between city zoning conventions and their influence on bike share usage. By utilizing the bike-share rides system dataset made available to the public and researchers by Bay Wheels Inc., who operate in San Francisco city, this paper uses a descriptive, predictive, and prescriptive analytics approach to assess the bike-share usage difference between and within the city zoning conventions. Four specific zones are considered in this study, along with weather events like average temperature and precipitation as covariates. Using data from the most recent, post-COVID two years of operations (2022-2023), the study finds support for the influence of city zoning conventions on bike share usage. The findings from between-zone differences and within-zone differences are discussed as these results set an optimistic tone for bike-share user preferences across city zones as a significant causal driver for bike-share operations design for effectiveness.

AI Technologies in Supply Chain: Aligning Innovation with Market Demand

mahsa soltaninejad, Oliver Schaer, Seung-Lae Kim

539rexel, Philadelphia, pen, USA

Abstract

The impact of consumer-facing AI tools within retail supply chains is significant, specifically aiming to prioritize these technologies based on the insights of both consumers and experts. While direct interaction between end consumers and the underlying AI technologies is minimal, the indirect feedback they provide through their usage patterns and preferences is invaluable. We explore which AI applications are deemed most beneficial by consumers for enhancing their shopping experience and compare these findings to expert opinions on which AI functionalities are most effective for optimizing supply chain operations. This analysis aims to pinpoint alignments and discrepancies between consumer preferences and expert opinions on AI, focusing on optimizing AI investments to balance market demands with operational efficiencies.

The research gathers expert opinions on the use of AI in supply chain management through detailed questionnaires. This questionnaire helps us understand the current use of AI in the supply chain. Additionally, we utilize Google Trends to assess the general interest and trends related to AI tools in the supply chain. Meanwhile, natural language processing (NLP) techniques analyze data from Reddit to gauge consumer sentiment about these AI tools. In the literature, we typically refer to the supply chain as the primary area of application, and similarly, we discuss consumer sentiment as a unified concept. This study identifies the AI features most valued by both experts and consumers, highlighting the alignment and discrepancies in their perspectives.

This study hypothesizes that advanced AI tools with strong customization and predictive capabilities are highly valued in supply chains. We investigate how varied assessments of AI functionalities, and the potential of generative AI can refine strategies, enhance operational efficiency, and improve customer satisfaction—key factors for maintaining competitiveness in dynamic markets.

Key words:

Consumer-facing AI tools, Supply chain optimization, Operational efficiency

Managerial Dysfunctions and Exogenous Shocks: The Case of the Hotel Industry During the COVID-19 Pandemic

William Kline, Richard Brown

Pennsylvania State University Harrisburg, Middletown, PA, USA

Abstract

Corporate managers are tasked with anticipating and addressing exogenous shocks, and their preparedness for such events is a concern for the firm's stakeholders. In this paper, we posit that managers that fail to adequately scan their external environment are prone to dysfunctional decision-making processes. Empirically, we provide a historical case study of top management teams in the Hotel and Lodging Industry to document the attention that these teams give to health-related exogenous shocks. The argument is that the SARS and Swine Flu epidemics were "dry runs" that should have led firms to enact buffers in the case of a subsequent pandemic. The subsequent pandemic happened in early 2020 as COVID-19 spread globally, leading to travel restrictions, lockdowns, and economic malaise. Our findings, however, point to scant attention of this type of risk due to managerial dysfunctions that hindered proper environmental scanning and effective decision-making.

Learning and Performance: An Investigation

Parvinder Gupta

Indian Institute of Management Ahmedabad, Ahmedabad, Gujarat, India

Abstract

Achieving excellent performance depends on a variety of determinants. The required skills and the knowledge base are considered among the most important ones. The ever-changing environment continues to pose challenges and opportunities demanding acquisition of different skill sets and knowledge. Nowadays, for example, huge curiosity especially among working professionals and researchers has emerged to explore deeply the world around Artificial Intelligence (AI). It rightly highlights the importance of ongoing learning; the need to continuously learn newer techniques, technology, skills, knowledge, and perspectives. In fact, learning how to learn has become critical to facilitate learning. The present paper, elaborating various theories and frameworks, examines how people can learn to be better learners, drawing implications for organizations to create climate supportive of learning, thereby inspiring outstanding accomplishments.

A Political Analysis Framework for Examining Power Dynamics in Open Strategy

Henry Adobor

Quinnipiac University, Hamden, CT, USA

Abstract

Open strategy is an emerging paradigm within the strategy-as-practice perspective. Open strategy is directed at leveraging the knowledge base of internal and external stakeholders, thereby expanding strategy formation beyond the C-suite. This research explores the intricate interplay of power dynamics within open strategy processes in organizations. The increasing emphasis on inclusivity in strategic decision-making challenges established hierarchies, often creating opportunities for power redistribution among organizational actors. The study underscores that open strategy may be inherently political, as individuals and groups leverage their power and influence to safeguard or advance their interests. This can manifest through strategic discourse, control over resources, and shaping decision-making processes. Such dynamics reveal the dual nature of open strategy processes: they can either promote collaborative engagement or entrench existing power imbalances. We explore how organizations can navigate these dynamics by fostering negotiation and dialogue among stakeholders and highlight the importance of creating mechanisms that enable previously excluded internal and external actors to contribute constructively, thus ensuring that inclusivity leads to innovative and sustainable outcomes through the co-creation of strategy and its implementation. By addressing the challenges of power and influence, the study provides insights into building equitable and effective open strategy frameworks. We offer the research and practice implications of the paper.

Filling in the Gaps: Missing Data Methods for Hofstede's LTO Cultural Dimension

Eugene Hahn¹, Jonathan Doh²

¹Salisbury University, Salisbury, MD, USA. ²Villanova University, Villanova, PA, USA

Abstract

Missing data is a pervasive problem in business research. Endpoint-heterogenous beta regression is a new method for modeling $y \in [0,1]$ data that has attractive characteristics. We apply this methodology to Hofstede's Long Term Orientation scale. We find we can estimate the distributions of the missing values. These can be used to reduce missing data problems in cross-cultural research.

The Effect of CEO and CFO Compensation Structures on the Quarterly Guidance Cessation Decision

William Kline¹, Tom Adams², Richard Brown¹

¹Pennsylvania State University Harrisburg, Middletown, PA, USA. ²La Salle University, Philadelphia, PA, USA

Abstract

This study examines the relationship between executive compensation and firms' decisions to cease issuing quarterly guidance. Building on prior research, this paper investigates whether a shift in compensation incentives toward long-term goals aligns with executives' claims of refocusing on long-term strategic objectives. Using a sample of 1,302 guidance stoppers and 756 maintainers between 2003 and 2012, we find that CFO compensation structures, favoring long-term over short-term incentives, significantly correlate with guidance cessation, supporting the long-term focus rationale.

Employee involvement in organizational social responsiveness initiatives; balance between internal and external empathy

Shanggeun Rhee, Glynnis Tan

Kean University, Union, NJ, USA

Abstract

Gaining employee engagement in organizational social responsiveness initiatives such as ESG, community outreach, and environmental sustainability is challenging. The prevailing top-down approach where initiatives are driven by leadership without strong employee involvement, leads to resistance from employees. Conventional solutions focus on managing the activities themselves such as program contents and design, communication and incentives but these approaches fail to foster genuine employee engagement. Drawing on empathy theory, this study argues that employee involvement in organizational social responsiveness initiatives is linked to their attachment to the organization. More specifically, when an organization balances its internal empathy with its external empathy, employees feel more connected and motivated to support these external initiatives. The key to gaining employee support for organizational social responsiveness lies in creating a balance between caring for employees internally and demonstrating empathy toward the external stakeholders, which in turn fosters employee attachment to the organization and encourages involvement in external social responsiveness initiatives.

From First Contact to Fatigue: The Hidden Toll of College Admissions Work.

RICHELYN PENN MEKILE, David Gefen

Drexel University, Philadelphia, PA, USA

Abstract

U.S. college admission counselors travel extensively to promote their university in high schools throughout their recruitment area, placing tremendous pressure on them and resulting in high turnover that negatively impacts student enrollment at a time when universities can hardly afford it. While work fatigue has been empirically studied across industries, including healthcare and blue-collar environments, no previous studies have examined the effects of work fatigue among college admission counselors. Addressing that question, this study examines survey responses from 192 counselors nationwide at traditional four-year higher education institutions through the theoretical lenses of Job Demands and Resources theory (JD-R) and measures self-reported fatigue levels with specific job demands. Initial findings show that physical fatigue is reduced by trust in the manager and satisfaction with compensation and is unrelated statistically to travel or workload. However, trust in the manager itself, while increased by satisfaction with compensation, is reduced by having to work an extended workday. The insights gained from this research can provide evidence-based solutions for senior leadership and human resource professionals at colleges and universities on the type of resiliency training needed to support the health and well-being of college admission counselors.

The Evolution of Teleworking: Analyzing Critical Success Factors for Sustained Excellence

Kellyann Kowalski¹, Jennifer Swanson²

¹University of Massachusetts Dartmouth, North Dartmouth, MA, USA. ²Stonehill College, South Easton, MA, USA

Abstract

Teleworking has rapidly evolved from a seldom used work arrangement to a mainstream practice, particularly accelerated by the global COVID-19 pandemic. This paper explores the historical development and current trends in teleworking by tracing its evolution. We begin with the early technological innovations in the 1970s, when remote work was a rarity, then move through telework's increasing adoption in the 2000s as internet access and digital tools enabled more flexible work arrangements. The next stage examines how the COVID-19 pandemic served as a spark for widespread remote work adoption, permanently altering workplace norms and expectations. Lastly, we look at post pandemic telework trends including hybrid models and new technological advances.

As we analyze the different stages of telework, we do so using a framework developed from previous research (Kowalski and Swanson, 2015) that identified critical success factors at the organizational, managerial, and employee levels. As remote work has become more common, its potential to increase work-life balance and autonomy has been cited as important factors related to job satisfaction. However, the challenge of maintaining team cohesion and collaboration has emerged as a critical concern, therefore many employers have rescinded prior telework arrangements. We discuss the importance and implications of critical success factors necessary for telework success, which include communication, support, and trust.

In conclusion, this paper offers insights into how teleworking is reshaping the future of work, highlighting both the opportunities and challenges it presents for employers, employees, and policymakers in a post-pandemic world.

A comparative look into coping and other organizational theories to explain
organizational resilience

Raktim Pal¹, Nezih Altay²

¹James Madison University, Harrisonburg, VA, USA. ²DePaul University, Chicago, IL, USA

Abstract

We extend coping, a theory adapted from psychology to supply chain disruption management, to organizational resilience and compare it to a select set of plausible organizational theories identified from the literature. We aim to answer two research questions: 1) How does coping explain organizational resilience? 2) How useful is coping theory in explaining resilience compared to other well-accepted organizational theories? We review relevant literature on disruption management and resilience focusing on organizational theories that have been used to shed light on these concepts in the Operations and Supply Chain Management (OSCM) field. We identify elements of resilience from the literature, and subsequently provide a comparative analysis of selected organizational theories through the lens of these elements of resilience. We also use the known-unknown framework to draw distinction of coping theory in explaining resilience. We argue that coping enhanced by continuous organizational learning and experience leads to resilience. Our comparative study shows that coping theory is a feasible and sound approach to explain the process of resilience. It performs better than other plausible organizational theories, such as resource advantage theory (RAT) and real options theory (ROT). OSCM literature is rich in papers that cover what and whys of resilience, but research explaining how organizations could build resilience is scarce. This paper fills this gap.

AI and supply chain resilience: benefits, challenges, and future

Pedro Reyes¹, Christopher Meyer¹, John Visich², Pat Jaska³

¹Baylor University, Waco, Texas, USA. ²Bryant University, Smithfield, RI, USA. ³University of Mary Hardin-Baylor, Belton, Texas, USA

Abstract

AI has garnered great interest in the supply chain from researchers and practitioners alike. It is rapidly reshaping the technological landscape and becoming an important component in the supply chain work workforce process. Likewise, it is being extended into the decision-making in many domains across supply chains. Gartner's hype cycle on generative AI places it positioned at the peak of inflated expectations, and it is projected to reach the transformational plateau of productivity benefit within two to five years. This paper reports perceived benefits, challenges, and future directions for building supply chain resilience based on multiple case studies and interviews with industry leaders on the needed skills for the “now” future workforce of supply chain professionals.

Generative AI in Supply Chain: A Systematic Review of Opportunities, Benefits, and Challenges

Musa Shaikh, Chandra Prakash, Pavan Mutha

University of the Cumberlands, Williamsburg, KY, USA

Abstract

This systematic review investigates the impact of Generative Artificial Intelligence (GenAI) on supply chain management (SCM). Though there is an acknowledgment of GenAI's potential in SCM, there remains a lack of research that thoroughly examines the opportunities, tangible advantages, and challenges linked to GenAI adoption. This literature review paper seeks to analyze the role of GenAI in SCM and fulfill three primary objectives. First, it identifies essential opportunities for GenAI implementation, encompassing its applications in demand forecasting, inventory control, and decision-making support systems. Second, it assesses the tangible benefits observed from current implementations, such as cost savings and enhanced operational efficiency. Third, the study investigates the challenges and obstacles to effective GenAI adoption, focusing on technical limitations, data quality concerns, workforce skill deficiencies, and ethical issues. This research employed a systematic review methodology, reviewing and synthesizing information from 24 peer-reviewed articles published between 2023 and 2025. The findings reveal substantial potential for GenAI in areas like demand forecasting, inventory optimization, and risk management while also emphasizing significant challenges related to data quality, integration, and organizational preparedness. The study offers two practical implications: a) the findings provide actionable insights for supply chain professionals, highlighting how GenAI can be integrated into SCM; b) the research underscores the necessity for strong data governance, ethical AI usage policies, and adherence to evolving regulations.

Data-Driven Supply Chains and Firm Performance: Unveiling the Role of Supply Chain Integration in Emerging Economies

Raza Rafique

Kean University, Union, NJ, USA

Abstract

This study explores the impact of data-driven supply chains (DDSC) on firm performance (FP) within an emerging economy context, utilizing a combined PLS-SEM and Artificial Neural Network (ANN) approach. Data were collected from 174 respondents responsible for managing supply chain practices in manufacturing firms. Drawing on the Resource-Based View (RBV), the research examines the relationships between DDSC, internal integration (II), customer integration (CI), supplier integration (SI), and FP. The findings indicate significant positive effects of DDSC on II, SI, CI, and FP, with SI mediating the role between DDSC and FP. Integrating PLS-SEM with ANN enables capturing both linear and non-linear relationships, providing a more comprehensive understanding of the complex dynamics within data-driven supply chains. These insights are valuable for firms that optimize supply chain integration and improve performance in emerging economies.

Trade credit contracts under weather risk

Piyal Sarkar¹, Mohamed Wahab Mohamed Ismail², Liping Fang²

¹Penn State University, Harrisburg, USA. ²Toronto Metropolitan University,
Toronto, Canada

Abstract

Trade credit is one of the widely used financing schemes today. Trade credit acts as an incentive to encourage buyers to increase ordering quantities as it allows delayed payments. The coordination mechanism under trade credit has become a hot research topic. However, no study to date has explored the role of trade credit financing in the case of supply chains under weather risk, even though the profitability of many industries is impacted due to weather. Results show that the trade credit contract with a weather rebate and financial hedging performs the best among the three contracts. This is a new idea in the area of trade credit financing. Our study contributes to the understanding of trade credit contracts and their implementations to withstand weather and credit risks in supply chains. It explores how weather rebate and financial hedging strategies can be implemented to enhance the performance of trade credit contracts under weather risk. Industries can use the contracts to coordinate supply chains under weather risk. Our analysis further shows that the designed trade credit contracts can be implemented under any market risks, which can be correlated or independent to the weather risk. This further generalizes the applications of the contracts under different risk scenarios.

Structuring Resilient Supply Chains: A New Blueprint for Network Design

Afrooz Moatari-Kazerouni, Yvonne Lederer Antonucci

Widener University, Chester, PA, USA

Abstract

In an era of heightened economic volatility and complex global interdependencies, designing resilient supply chain networks has become vital for organizational success. This study investigates the resilience strategies applied within three prominent supply chain network design (SCND) structures: centralized, decentralized, and distributed. Using data from a comprehensive globally-distributed survey of over 900 supply chain professionals, the research utilizes quantitative analysis to explore the relationships between these SCND structures and a framework of eight resilience strategies and eighteen critical capabilities.

The results indicate that distributed supply chains are strongly associated with resilience, demonstrating superior performance across fourteen key measures. These networks, characterized by their flexibility and decentralized decision-making, enable organizations to respond to disruptions with greater speed, adaptability, and visibility, offering a more robust defense against uncertainty compared to centralized systems. Distributed supply chains also facilitate better collaboration, strategic alignment, and proactive planning, making them more effective at managing risks. Additionally, this research identifies two main clusters of resilience capabilities, i.e., proactive and reactive, which align with existing theoretical models in literature.

The study fills a crucial gap by providing empirical evidence on how SCND structures influence resilience, offering actionable insights for organizations seeking to enhance their supply chain risk management and sustain operational continuity. The findings underscore the importance of designing resilient supply chains that are adaptable and robust in the face of uncertainty, reinforcing the critical need for distributed network models in global supply chain strategies. By offering data-driven insights, this research contributes to the broader understanding of how strategic supply chain design can bolster resilience and improve long-term performance in turbulent environments.

Indoor Farm Supply Chain Management

Eric W Stein^{1,2}, Afrooz Moatari-Kazerouni³

¹Penn State, Malvern, PA, USA. ²Center of Excellence for Indoor Agriculture, Philadelphia, PA, USA. ³Widener University, Chester, PA, USA

Abstract

This study examines the complexities of supply chain management within the context of indoor farming, also known as controlled environment agriculture (CEA). As global food security concerns and sustainability challenges increase, the optimization of processes in controlled agriculture production environments has become increasingly critical. Indoor farming represents a significant innovation for agriculture by offering the potential to produce food in urban as well as rural environments while reducing the use of land, water, and the negative impacts of farming on the environment (Stein 2021).

This study addresses the critical needs for robust supply chain frameworks to support the scalability and economic viability of indoor farming. The application of the Supply Chain Operations Reference (SCOR) Model can provide a structured means to assess and enhance operational efficiency, reliability, and responsiveness in these systems. Moreover, while existing literature on supply chain management in traditional agriculture is extensive, there is a notable gap in research specifically focused on indoor farming supply chains. We contribute to filling this gap by applying the SCOR Model to indoor farming by providing a detailed analysis of process flows, automation opportunities, and performance metrics.

The findings of this research have significant practical implications. By identifying key areas for process improvement and automation, the study offers actionable strategies to enhance productivity and to reduce operational costs. Facility layout recommendations, along with a focus on automation and integration, provide a blueprint for designing efficient indoor farming operations. Additionally, the performance metrics derived from the SCOR Model offer a benchmark for continuous improvement and help to navigate the economic challenges of this emerging industry.

Collaborative Climates and Supply Chain Flexibility: Keys to Competitive Advantage

Yan Jin

Quinnipiac University, Hamden, CT, USA

Abstract

The continuous interruptions and unprecedented events in the global supply chain highlight the urgent need to build resilient supply chains (Roshani et al., 2024). To achieve sustainable competitive advantages in such a competitive environment, a manufacturing firm must develop proper flexibility in its suppliers and within the company itself (Narasimhan et al., 2004; Perez-Perez et al., 2021). Although flexibility has been extensively studied for decades (Slack, 1987; Stevenson and Spring, 2007), evolving challenges, including geopolitical tensions and labor strikes, have underscored the pivotal role of supplier flexibility, distribution flexibility, and supplier management flexibility in determining a manufacturing firm's competitive advantages (Zhang et al., 2006). These flexibilities cannot be created in a vacuum but are instead rooted in a collaborative climate shared by both suppliers and the manufacturing firm, which is characterized by open communication, trust, and long-term partnerships (Das & Teng, 1998; Huang et al., 2020; Narayanan et al., 2015). Ultimately, this collaborative climate leads to sustained competitive performance by enhancing the overall resilience and adaptability of the supply chain.

Despite the increasing recognition of the importance of collaboration and flexibility, there is a lack of research exploring these factors in both manufacturers and their suppliers. This study addresses this gap by examining the interplay of collaboration and flexibility in a unified model encompassing both parties. A large-scale survey was conducted, collecting 201 responses from senior manufacturing executives. Confirmatory Factor Analysis (CFA) will be used to evaluate the reliability and validity of each construct, and Structural Equation Modeling (SEM) will be used to test the structural fit (Fornell & Larcker, 1981; Shah & Goldstein, 2006). The findings will demonstrate how collaborative and flexible practices translate into competitive advantages, such as superior quality, dependability, and time to market, posing significant managerial implications for both manufacturing companies and their suppliers.

Resilience to Supply Chain Disruptions: An Investigation of the Roles of Innovativeness and Green Supply Chain Management Practices During the COVID-19 Pandemic

Lindsay Givens

Southern Arkansas University, Magnolia, Arkansas, US

Abstract

The number of global supply chain disruptions is increasing at an alarming rate. These disruptions vary in terms of source, duration, and scope, but all pose major threats to the continuity of firms worldwide. Because of this increase in turbulence, there has been heightened interest in recent years surrounding the development of firm-level resilience to such disruptions. The supply chain sustainability research stream is relatively well-developed; less is known, however, about specific strategies for strengthening resilience. This study aims to contribute to the resilience literature by examining the viability of innovativeness and the adoption of green supply chain management practices as supply chain resilience enhancement strategies. Drawing from dynamic capabilities view, this study examines archival data from The World Bank through the lens of the COVID-19 pandemic disruption. Using propensity score matching, genetic matching, and random forest classification, this study highlights the importance of adopting green practices as a means of mitigating the damages incurred by disruptive events. This study advances the theoretical understanding of supply chain resilience and offers practical applications for enhancing resilience at the firm level.

Cross-channel effect in omnichannel retailing

Shengbin Wang

The College of New Jersey, Ewing, NJ, USA

Abstract

This research studies the cross channel effect in an omnichannel retailing setting. A benchmark game model along with two other comparative models that involve online and offline retailers are developed. We analyze how online return rates and costs impact each model's performance and optimal strategies. We then identify conditions under which the online retailer adopts an omnichannel strategy and assess its effect on the offline retailer's profits. Comparative analysis shows that the online retailer prefers the omnichannel strategy when certain thresholds are high, with the choice between the two comparative models driven by return processing costs. Post-omnichannel implementation, the offline retailer experiences reduced profits on initial products but gains cross-selling opportunities. Lastly, how coupons from the offline retailer influence the online retailer's omnichannel strategy is explored.

Enhancing Urban Logistics Resilience with Parcel Locker-Assisted Dual-Mode Multi-Objective Vehicle Routing

Dan Liu

Kean University, Union, New Jersey, USA

Abstract

Urban logistics systems are increasingly challenged by rising demand, unpredictable disruptions, and the need for sustainable operations. Ensuring resilience in these systems has become a critical focus as cities strive to maintain efficient and reliable distribution networks under uncertain conditions. This study addresses the critical need for resilience in urban logistics by optimizing distribution paths through the integration of parcel lockers under stochastic demand and delivery times. A novel framework, Multi-Objective Dual-Mode Distribution Vehicle Routing Problem with Parcel Lockers (MODD-VRP-PL), is proposed, incorporating a composite objective function that balances efficiency and resilience. By advancing beyond traditional single-objective frameworks, this research emphasizes adaptability and robustness in logistics operations.

To advising the efficiency of current algorithm for solving the multi-objective VRP, we developed a “Random-Optimization” based Heuristic Algorithm (ROHA) to consistently identify efficient and resilient paths under variable conditions. To evaluate the role of parcel lockers in enhancing resilience, this research employs sensitivity analysis to simulate various disruption scenarios and demand fluctuations. The sensitivity analysis is designed to assess the system’s performance across multiple dimensions, including delivery efficiency, and stability under adverse conditions. By varying key parameters, such as demand uncertainty, and the availability of lockers, the analysis provides insights into how well the logistics network adapts and maintains service levels.

This study underscores the potential of MODD-VRP-PL and ROHA to bolster the resilience of urban logistics systems, ensuring operational efficiency even in the face of uncertainty and disruptions.

Deterring counterfeits through the introduction of a store-brand product

Jing Chen¹, Bintong Chen²

¹Dalhousie University, Halifax, NS, Canada. ²University of Delaware, Newark, DE, USA

Abstract

Online retail enhances convenience but also exposes consumers to counterfeits. This study explores how the introduction of a store-brand product by an e-commerce platform can reduce counterfeit sales when a retailer sells a manufacturer's authentic product with counterfeits. The effectiveness depends on the authentic product's production cost and the store-brand product's quality. The introduction of the store-brand product can deter counterfeits when the store brand is low quality with moderate authentic product costs or high quality with higher costs. However, the platform may avoid introducing a store brand despite its deterrent effect. Higher production costs increase the likelihood of the introduction of a store-brand product, which may increase the retailer's demand.

Supply Chain Contracts and Quality Investment with Blockchain Adoption

Min Wang

Drexel University, Philadelphia, PA, USA

Abstract

Blockchain technology enables traceability of products in the supply chain and allows companies to gain access to key information such as product source, transportation environment and quality. The study explores the impact of enhanced information transparency as the result of adoption of blockchain technology on supply chain contracts and quality investment.

Collaborative New Product Development: The Implications of Long-term vs.
Short-term Contract Choice on Cost-sharing in Vertical Channels

Abhishek Roy¹, Ganesh Balasubramanian²

¹Temple University, Fox School of Business, Philadelphia, PA, USA. ²T A Pai
Management Institute, Manipal, KA, India

Abstract

Problem Definition: It is well known that a buyer expects price commitment from an upstream supplier when they collaborate to improve product quality. This insight is based on the notion that an upfront price commitment from the supplier allows the buyer to confidently co-invest without having to worry about the supplier exploiting the buyer with a future price rise. Existing research on supplier-buyer co-creation has not addressed the impact of the buyer's ability to influence the supplier's demand in multi-period selling environments by carrying strategic inventory. We investigate how the possibility of strategic inventory influences the buyer's co-creation effort, product quality, and the supplier's contract choices between short-term transactional and long-term commitment contracts.

Methodology/Results: Using a stylized game-theoretic model, we consider the collaborative interactions between a supplier and a buyer over a two-period horizon in which demand is deterministic. Our analysis reveals that the buyer's collaboration effort decreases with price commitment for sufficiently low values of holding cost. Further, we find that the supplier might prefer a short-term contract even when it attracts a lower quality cost-sharing effort from the buyer compared to the long-term contract.

Managerial Implications: Considering the impact of short-term dynamic contracts on supplier-buyer collaboration efforts in a multi-period setting, we contribute to the literature on collaborative product development and strategic contracting. Our research enables suppliers and buyers who engage in collaborative product development to optimally determine their efforts and contract choices.

Combating Greenwashing in Green Supply Chains: A Game-Theoretic Approach to Auditing Practices with Blockchain Technology

Shubham Gupta, Vignesh Kesavaprakash

New Jersey Institute of Technology, Newark, NJ, USA

Abstract

As consumer demand for sustainable practices continues to grow, firms face increasing pressure to adopt transparent supply chain operations. However, the prevalence of greenwashing undermines consumer trust and hinders progress toward genuine sustainability. This research investigates the strategic interactions between two competing firms in a green supply chain context using a game-theoretic framework. We model scenarios where one firm adopts rigorous auditing practices, such as Blockchain Technology (BCT), while the other opts out, examining how these decisions influence brand image, consumer trust, and market outcomes. The study employs game theory to analyze the conditions under which BCT adoption effectively mitigates greenwashing, highlighting the trade-offs between the costs of implementation and the benefits of increased transparency. Our findings suggest that consumer awareness and the risk of greenwashing exposure significantly impact firms' strategic decisions, emphasizing the role of advanced auditing technologies in fostering sustainable supply chain practices. This work contributes to the intersection of sustainability management and decision sciences by providing actionable insights for firms navigating the complex dynamics of consumer expectations and competitive pressures.

Impact of cocreation under supply chain encroachment setting

Abhishek Roy¹, Piyal Sarkar², Shubham Gupta³

¹Temple University, Philadelphia, USA. ²Penn State University, Harrisburg, USA. ³New Jersey Institute of Technology, Newark, USA

Abstract

The increasing complexity of supply chains has led firms to redesign supply chain strategies based on the collaborative value created through cocreation with other supply chain members. Cocreation is an economic strategy that binds supply chain members to jointly improve the output. Previous literature has investigated the impact of cocreation under different settings, but to the best of our knowledge, the impact of cocreation under supply chain encroachment setting has not been investigated so far. Through this research, we try to address this major gap in the literature. Supplier encroachment mitigates double marginalization and secure Pareto improvements for both parties. Our research investigates on how cocreation can add value under supplier encroachment environment for a two party supply chain. The study provides several managerial insights by integrating the cocreation and encroachment strategies.

The Role of Generative AI in World-Class Logistics and Supply Chain Management

Angappa Gunasekaran

Penn State University, Middletown, PA, USA

Abstract

This paper explores the application of Generative AI in world-class logistics and supply chain management (WCLSCM). It begins by examining the role of Generative AI in advancing logistics and supply chain development and operations. Next, it introduces a framework for achieving WCLSCM. Building on this framework, the paper explores how Generative AI can support strategic and operational decision-making to enhance logistics and supply chain performance. Finally, it highlights the opportunities and challenges of implementing Generative AI in WCLSCM.

Regulating safer consumer products through supply chain mapping

Penina Orenstein¹, Elena Galkina²

¹Seton Hall University, South Orange, NJ, USA. ²California Department of Toxic Substances Control, California, USA

Abstract

This research, conducted in partnership with the Safer Consumer Products program (SCP) of California Department of Toxic Substances Control, aims to enhance the implementation of SCP regulations through analysis of supply chain maps. To achieve this, we map supply chains of two key players in cosmetic product manufacturing, which fall under SCP's current Priority Products Work Plan. Using Bloomberg financial data, we compiled and organized financial metrics for these two cosmetic manufacturers to facilitate the supply chain mapping and explain the strength of the relationships between the companies, which can become responsible entities. The resulting supply chain maps will inform SCP's decision-making process regarding product and chemical prioritization, approach to alternatives analysis, and compliance.

Multi-tier supplier selection using total cost of ownership and data envelopment analysis

Feras Saleh, Anjali Awasthi

Concordia University, Montreal, Quebec, Canada

Abstract

Quality management across multiple tiers is vital to minimize cost of quality in global supply chains. In this paper, we are addressing the problem of supplier selection in multi-tier global supply chains with the purpose of overall quality management. A hybrid approach based on Total Cost of Ownership (TCO) and network Data Envelopment Analysis (DEA) is proposed. The TCO looks beyond the quoted cost to cover additional true costs related to the entire purchasing cycle. The cost categories included are quoted price, manufacturing costs, quality costs, design costs, logistics costs, after sales service, and social/environmental costs. Network DEA is used to rank the suppliers based on the TCO cost categories. The advantage of network DEA is its ability to investigate intermediate linkages between different stages of the supply chain. The results of network DEA are efficient suppliers and improvement targets for inefficient suppliers for improving overall quality in global supply chains. A numerical application is provided. Comparison of the approach with some of the existing approaches available in literature is done using a case of Iranian beverage corporations.

An Interdisciplinary Approach to The Synthesis of a Novel Circular Value Chain

Eric Addeo¹, Deborah Helman¹, David Walters²

¹DeVry University, Iselin, NJ, USA. ²Independent Researcher, Port Macquarie, Australia

Abstract

When it comes to the topic of a “Value Chain,” there is a transformative and empowering vision being pursued by business leaders and technologists of an evolutionary and powerful value chain architecture that leverages off recent advances in artificial intelligence, block chain mechanisms, and the metaverse to create a new value chain architecture. This new architecture also has a feedback channel – a circular feedback control model that supports green innovation through minimization of environmental waste. This new system inherently supports opportunities for a dynamic, responsive, and resilient end-to-end value chain system that is particularly well- matched to the realities of globally changing markets, new centers of production efficiency, and the evolution of customer demand. This architectural approach is believed to be novel, and it builds on the previous work of the authors during the period from 2010 to the present. This new architecture has the potential to realize new business synergies that promise significantly improved efficiencies amongst business partners through real-time and dynamically adaptive production resources that are matched to the reality of a rapidly changing source environment with simultaneous changes in end-user preferences. The focus of this contribution will center on the technical landscape, where a brief perspective on the evolution of enabling technologies will be given that directly support the notion that many of the key technological underpinnings needed to realize a vision of a circular value chain are already in place! For example, with the emergence of globalization and digitalization, AI has gained attention to connect businesses globally by improving supply chain processes like production, raw material inspection, and distribution. This response is so well matched to volatile industries where it is always challenging to quickly respond to changes in the breadth of trends and continuously evolving demands of customers.

Plastic Recycling Fueled with Blockchain-Driven Tokenization: ESG Optimization and Implications

Aichih Chang¹, Nesreen El-Rayes², Jim Shi¹, Fuqin Zhou¹

¹New Jersey Institute of Technology, Newark, NJ, USA. ²Georgia State University, Atlanta, GA, USA

Abstract

The plastic waste crisis has been exacerbating nowadays. To address the burgeoning challenge, one innovative solution is to incentivize recycling via leveraging tokenization thanks to Blockchain technology (BCT). In this study, we devise an optimization model considering the tokenization to examine the environmental and social governance (ESG) performance. In particular, we investigate the optimal decisions of tokenization rewards and the government enforcement as well as their interplay on the ecosystem. We demonstrate the potential for significant ESG improvements by managing enforcement and tokenization jointly. Extensive numerical experiments and sensitivity analysis are performed to provide rich insights. For example, the value of tokenization in terms of ESG is visualized.

Exploring the Sustainable Development of Taiwan's Taoist Temple Culture through the System Dynamics Approach

Tung-Shan Liao, Shan-Hai Lin Lung

College of Management, Yuan Ze University, Taoyuan, Taiwan

Abstract

This study utilizes a system dynamics approach to construct a causal feedback loop model to promote the sustainable development of Taiwan's temple culture. The model underscores the vital role of temples and folk beliefs in fostering regional cultural sustainability, strengthening community identity, and building social consensus.

As deeply embedded cultural phenomena, temples function as both sanctuaries for spiritual solace and tangible representations of community cohesion and collective faith. However, modernization—driven by industrialization, urbanization, and globalization—poses significant challenges to these traditions. Secularization and socio-economic shifts have gradually diminished younger generations' connection to folk beliefs, threatening the continuity of traditional temple culture.

Innovative adaptations in religious practices, integrating cultural marketing with community-oriented activities, offer promising avenues for revitalizing these traditions. Temples, by hosting religious festivals and charitable events, not only solidify their follower base but also engage younger participants, narrowing intergenerational cultural gaps and enhancing the transmission of heritage. Additionally, modernized practices such as “religious governance” illustrate the dynamic evolution and adaptability of folk beliefs in addressing contemporary social changes.

Policy and educational interventions are critical to mitigating these challenges. Systematic cultural promotion and embedding temple culture within the broader framework of sustainable development can enhance cultural identity and safeguard intangible cultural assets. This study emphasizes that balancing preservation with innovation is pivotal to ensuring that Taiwan's rich religious and cultural heritage continues to thrive in an increasingly modernized society.

AI and RBV: Generating Resilient Supply Chains

Mark Lennon¹, Jerry Carbo², Ian Langella²

¹Pennwest University, California, PA, USA. ²Shippensburg University,
Shippensburg, PA, USA

Abstract

As global supply chains grow more complex, incorporating AI with responsible business practices becomes essential. Applying RBV theory, this paper explores how AI, as a valuable organizational resource, offers a competitive edge by optimizing logistics, demand forecasting, and HR operations. We examine both the benefits and challenges of AI, focusing on its role in mitigating risks and addressing sustainability concerns. Through real-world examples, we propose strategies for employing AI to boost efficiency, resilience, employee relations, and community well-being.

Judgmental forecasting for decarbonization pathways in the European energy system

Ali Gamaleldin Karam¹, Mohamed Elshahhat Ali Khedr¹, Francesco Cariat¹, Emanuele Giovanni Carlo Ogliari¹, DINA (KONSTANTIA) LITSIOU², Vasileios Bougioukos³

¹Politecnico di Milano, MILANO, Italy. ²MANCHESTER METROPOLITAN UNIVERSITY, MANCHESTER, United Kingdom. ³NOTRE DAME LONDON, LONDON, United Kingdom

Abstract

Judgmental forecasting is a qualitative method in the forecasting arena based on individuals' opinions rather than on quantitative data. It is predominantly used in the business and management sector. This study explores the application of the judgmental forecasting technique to predict decarbonization pathways in the European energy sector. Twenty energy engineering master's degree students participated under the supervision of an expert moderator. Following the brainwriting technique training, participants employed structured analogies and scenario development methods. In the second phase, a survey based on the opinions of 80 experts, and the results were analysed. Key insights highlighted nuclear energy, renewable integration, grid management, and electrification, alongside concerns about material availability and geopolitical dependencies.

Literature Review on Resource Efficiency and Value Chain Innovation in Semiconductor Industry

Xiangrong Liu¹, Samuel Serna-Otalvaro¹, Anu Agarwal², Pradnya Nagarkar², Lionel Kimerling², Ajay Gupta², Sajan Saini², Takahiro Ono², Beth Unger²

¹Bridgewater State University, Bridgewater, MA, USA. ²Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

As the demand for cloud computing and AI systems escalates, so too does the need for advanced semiconductor chips. This surge has intensified scrutiny of the environmental impact of existing technologies, while simultaneously driving innovation in more resource-efficient semiconductor technologies. This work will review relevant literature on this topic and provide insights into value chain innovations that can contribute to a more resource-efficient semiconductor industry – from raw materials, design, fabrication, packaging, testing, component use, and end of life. This work was funded by the NSF Convergence Accelerator Track I: FUTUR-IC: A Sustainable Microchip Manufacturing Alliance, Award Number ITE-2345076.

Keywords: Semiconductor Manufacturing, Value Chain Innovation, Resource Efficiency

A Practical Guide for Navigating the PTE Tax Landscape

Caitlin Sizemore, Alisha Harper

Bellarmino University, Louisville, KY, USA

Abstract

The State and Local Tax (SALT) deduction has historically allowed taxpayers to deduct the full amount of state and local income, property, or general sales tax paid as an itemized deduction on their federal tax return. However, with the passage of the Tax Cuts and Jobs Act (TCJA) in 2017, this deduction was capped at \$10,000, creating disparity for taxpayers in high tax states and causing those high tax states to look for workarounds. In response, many states have introduced the Pass-Through Entity (PTE) tax as a workaround to the SALT cap. Pass-through entities are typically defined by the flow through of income and deductions directly to their owners, who then pay any resulting tax at the owner level. The new PTE tax allows the entity to pay the tax at the state level, creating a credit or deduction that flows through to the owner.

This revolutionary change presents opportunities and challenges for taxpayers, offering benefits, but also creating uncertainty. This research examines these issues, using a state-by-state analysis of the PTE election. The benefits and uncertainties surrounding the PTE election will be further highlighted by a visualization of multiple tax returns, comparing high-income tax, low-income tax, and no income tax states pre-TCJA and post-TCJA. The findings offer a practical guide for taxpayers and practitioners navigating the PTE tax landscape, helping them make informed decisions in light of these changes.

Educational Growth and Student Access at the University of Southern Maine:
Expanding the Modalities and Frequency of Course Offerings

Isabella Rossborough, Simon Gabaree

University of Southern Maine, Portland, Maine, USA

Abstract

The overall objective of our study is to examine the University of Southern Maine [USM] students' satisfaction with course offerings. USM provides post-secondary education to more than 7500 students with more than 5500 undergraduate students, 1700 graduate students, and 260 law students (USM, 2022). USM spans across three campuses in Gorham, Portland, and Lewiston offering “nearly 300 majors, minors, certificates, and graduate degrees” to students (USM, 2024). USM caters to a wide variety of students, traditional and non-traditional, diverse age groups. In 2022, USM reported having a total of 2777 students pursuing education on a part-time basis and 1458 students on non-degree seeking tracks, meaning they are not pursuing a degree, but rather taking a course for personal development, professional growth, continuing education, etc. (USM, 2022 & 2024).

The problem facing students at USM is difficulty planning their schedule due to classes only being offered specific semesters, along with some classes only being offered through one modality or during a specific time of day. While it varies by program, some courses are only available to students one semester per academic year, meaning if the course fills up, does not fit into your schedule, or you have yet to complete the prerequisites you will have to wait a full academic year before being able to take that course. This has the potential to stall students' graduation timeline forcing them to put their post-graduation plans, such as working, and applying for graduate or research programs. On hold. The specific purpose of our study is twofold. First, we test the relationship between course availability (availability per semester, morning, afternoon, evening), amount of courses (in-person courses, online/hybrid courses), and student satisfaction with modalities and course frequency. Second, we test the relationship between student satisfaction with modality and course frequency, and education value.

Pixel Meets Algorithms; ARTificial Intelligence Revolutionizes Video Game Design

Devon Tolbert

York College of Pennsylvania, York, PA, USA

Abstract

This research investigates how artificial intelligence can be systematically integrated into every stage of game design, from initial concept to final implementation. The central research question is: “Can AI tools transform the game development workflow, creating a more efficient and innovative design process?” The study aims to demonstrate AI’s potential to revolutionize productivity and creativity in game development. For example, smaller experimental games like the classic Snake game have already been developed using generative AI and chatbots. “Echoes of the Abyss” (EOTA)—a 2D pixel adventure game that began development in August—serves as the primary research platform. The research methodology is structured across four phases: ideation and world-building, asset creation, animation, and code generation. Each phase is designed to leverage specific AI technologies like Stable Diffusion, ForgeUI, Claude 3.5, and GPT-4o, creating a holistic approach to game creation that accelerates development of complex character designs, narrative frameworks, and game specifications. Currently in its second phase of development and being created in Godot, EOTA is a larger experiment that combines a notable game engine with a wide variety of AI tools to test a completely new workflow. These tools significantly reduce traditional development timelines while maintaining high creative standards, challenging conventional development methodologies. The approach allows for faster operations and innovation in game design, showcasing AI’s potential as a collaborative creative tool. The experiment will continue throughout the rest of the year as the project enters into its later stages.

Digitizing historical logistics industrial maps and studying the growth and development of Jersey City Ports post world war

Ankith Anil Kadagadakai, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Abstract

This project digitizes old maps of Downtown Jersey City from the 1950s and 1960s using Blender and Geographic Information Systems (GIS). Despite their historical and educational significance, paper maps have certain drawbacks, such as being susceptible to physical harm, being challenging to update, and lacking interactive features, creating challenges for business innovation. These difficulties highlight the necessity of digitization to improve accessibility and preserve information. Through the integration of a literature review and a historical narrative, this study lays the groundwork for several applications. The paper discusses the advantages and disadvantages of Blender and GIS while outlining the processes for digitizing and visualizing maps. The project also includes creating 3D models of key landmarks in Blender, integrated into GIS, enhancing the visualization of the city's historical landscape, and facilitating a comprehensive analysis of its growth over the decades. This effort bridges the gap between past and present, fostering an appreciation for Jersey City's historical evolution.

High Scores, Dark Agendas: Radicalization Within Gaming Platforms

Nevin Ferry, Tamara Schwartz (mentor)

York College of Pennsylvania, York, PA, USA

Abstract

Extremist groups are widespread across various online platforms, and the online gaming industry is no exception. As this industry continues to grow, it attracts a diverse range of audiences, particularly young individuals. These young players are especially vulnerable, given their unrestricted access to online spaces. Extremists take advantage of these vulnerabilities, targeting youth to further their agendas. The real concern lies in understanding the specific strategies these groups use, how they are integrated into online gaming environments, and what efforts are currently in place to address these issues. As online gaming becomes an increasingly popular pastime, it is crucial to examine how these extremist activities are facilitated and what can be done to prevent the exploitation of young players. Combatting this problem requires a multi-faceted approach, including improved moderation, education, and awareness about the potential dangers of online radicalization in gaming communities.

Assessing the Security of Operational Technology Industrial Control Systems
Controlled by Information Technology Assets

Conner Everett, Tamara Schwartz (mentor)

York College of Pennsylvania, York, PA, USA

Abstract

Operational Technology and Information Technology are commonplace in the current business environment. While security and standardization of these technologies have been documented thoroughly the intersection has been left largely ignored, resulting in outdated and vulnerable systems. With a lack of documented overlap, solutions are few and far between. By using a thought experiment methodology, potential solutions have been evaluated and considered. These systems may be able to continue to serve their purpose while remaining secure through the use of virtualization, airgapping, and other potential solutions.

Patient Privacy: An Ethical Analysis of Emerging Tech in Hospitals

Ian Burns, Tamara Schwartz (mentor)

York College of Pennsylvania, York, PA, USA

Abstract

Hospitals are using robots, AI, and smart devices to improve patient care, but these come with risks. Cyberattacks could leak sensitive health information data or even cause harm to patients. AI can make decisions without doctors, raising concerns about mistakes and who's responsible. Devices like fitness trackers and pacemakers collect data, but it's unclear who owns it, patients, hospitals, or companies. Supply chain problems, like chip shortages, make this situation worse. This research looks at the risks and how to mitigate risks to keep hospitals safe and trustworthy while using these new technologies.

From Reps to Replicas: The Rise of Fitness Digital Twins

Zachary Deal, Tamara Schwartz (mentor)

York College of Pennsylvania, York, PA, USA

Abstract

The fitness industry increasingly collects extensive data on its users, ranging from personal identifiers to workout habits and device usage. The internet of Things, the rise of health and fitness apps, and the creation of digitally connected gym equipment result in comprehensive, virtual representations of users. Using a systematic review to gather data and a test matrix methodology, this research tests the hypothesis that a “digital twin” can be created from the integration and synthesis of data collected in fitness centers, health and fitness apps, and digital fitness equipment.

Investigating the Impact of Tariffs on Global Trade in the Trump 2.0 Era

Youngwoo Cho, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Abstract

With the beginning of Donald Trump 2.0, economic attention has turned to trade protectionism of the United States. The protectionism focuses on the high tariffs that Trump plans to impose following his term. Tariff increases primarily drive the ongoing geopolitical trade tensions with China. Additionally, Trump 2.0 announced that it would conclude and renegotiate the current U.S.-Mexico-Canada Agreement (USMCA) and the Korea-U.S. Free Trade Agreement. This study will analyze and visualize the trade volume between China and the U.S., the USMCA, and China and Mexico. Preparing measures for the swift and smooth conclusion of trade agreements between countries is crucial.

Reducing Fresh Produce Waste Using Cellulosic Bio-Coating

Joselyn Reyes, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Abstract

Fresh produce waste is a significant challenge in Lidl's operations, impacting profitability, environmental sustainability, and customer satisfaction. This proposal recommends introducing cellulosic bio-coating—a biodegradable and edible coating for fresh produce that extends shelf life while maintaining quality. The solution aligns with Lidl's commitment to reducing food waste and enhancing operational efficiency. By adopting this innovative approach, Lidl can reduce spoilage, improve customer satisfaction, and strengthen its reputation as a sustainable retailer.

Exploring the Town Management and University Partnership Impact on Community and Economic Development in Bridgewater, Massachusetts

David Alatorre

Bridgewater State University, Bridgewater, MA, USA

Abstract

This research investigates the Town of Bridgewater, Massachusetts with a mixed method. The interconnection between community and economic development was explored and analyzed through a qualitative approach to understand how these factors and their underlying elements impact the Town of Bridgewater stakeholders. As a key cornerstone in this study, the *Town of Bridgewater Master Plan, (2022)* provides a compendium of influential priorities for the Town regarding community and economic development, current efforts initiated to support those priorities, barriers hindering further advancement, and future opportunities. Additionally, the unique characteristic of Bridgewater being a university town is examined for its role in shaping community and economic growth. Moreover, feedback from small local businesses is also incorporated, emphasizing their perspectives on how the town can support their success and contribute to the broader community development. The analysis of these significant town viewpoints enables this report to provide a comprehensive overview of what is influential, the barriers that hinder it, and the further opportunities to advance community and economic development in the Town of Bridgewater.

Deep Reinforcement Learning for the Heterogeneous Capacitated Vehicle Routing Problem

Yuanyumeng Zhu¹, Dan Liu²

¹Department of Management, Kean University,, Union, New Jersey, USA. ²Department of Management, Kean University, Union, New Jersey, USA

Abstract

The Heterogeneous Capacitated Vehicle Routing Problem (HCVRP) is a significant area of research in vehicle routing optimization, particularly in supply chains with fluctuating demands and complex transportation conditions, such as fresh food delivery and pharmaceutical logistics. Effectively addressing HCVRP can improve delivery efficiency, reduce costs, lower carbon emissions, and enhance customer satisfaction. However, traditional optimization methods often lack the flexibility and computational efficiency needed to manage the dynamic and complex nature of HCVRP. This study presents a novel deep learning-based model that leverages modern optimization techniques to enhance routing efficiency and address complex dynamic constraints.

The study introduces two key innovations. First, a deep reinforcement learning (DRL) framework is developed to tackle dynamic demands and diverse constraints, such as vehicle capacities, speeds, and order priorities. By continuously interacting with a simulated environment, the model learns optimal routing strategies, improving adaptability to changing demands and conditions. A hierarchical DRL architecture further enhances performance by decomposing tasks into global route planning and local dynamic adjustments, reducing training and inference times. Second, the study integrates the Linformer attention mechanism with multi-relational modeling to improve computational efficiency for high-dimensional problems. By incorporating a multi-relational node selection decoder, the model effectively unifies constraints such as vehicle capacities, service time windows, order priorities, and dynamic demands. This approach captures intricate relationships within delivery networks, improving the reliability and practical applicability of routing solutions, particularly in dynamic and complex logistics scenarios.

This study validates the Deep Reinforcement Learning (DRL) framework for solving the Heterogeneous Capacitated Vehicle Routing Problem (HCVRP) using a simulated environment with fixed constraints. Training includes hierarchical global planning and local adjustments. The model is evaluated for efficiency, constraint satisfaction, and responsiveness. Analysis by removing key components highlights each innovation's contribution. The framework demonstrates practical value in logistics.

Information Gain-Based MCDM Approach for Urban Critical Freight Corridor Identification using Spatial-Temporal Data Analysis

YIYANG LIU, DAN LIU

KEAN UNIVERSITY, UNION, NEW JERSEY, USA

Abstract

Urban freight corridor optimization presents a multifaceted challenge in rapidly evolving cities, requiring a careful balance between economic efficiency and public safety. Existing approaches often fail to capture the complex interplay between these metrics, limiting their applicability in real-world scenarios. To address these gaps, this study proposes an information gain-based multi-criteria decision-making (MCDM) framework that integrates multimodal data, including regional GDP, delivery stop data, vehicle flow, traffic accidents, and points of interest (POI). The framework tackles the critical challenge of integrating diverse and often conflicting metrics by quantifying the relative importance of each variable through data-driven coefficients. By capturing nuanced interdependencies, this approach overcomes the limitations of existing methods that typically isolate economic or safety factors. Trained on historical data and validated on separate subsets, the model demonstrates scalability and adaptability, making it a robust tool for identifying critical corridors that are both economically vital and safety-sensitive across diverse urban contexts. A key innovation of this study lies in the development of the information gain-based MCDM framework, which not only incorporates economic and safety metrics as input variables but also uncovers their interdependencies. By identifying high-priority freight corridors that balance economic significance with safety concerns, the framework provides actionable insights for urban policymakers. These insights promote sustainable mobility, reduce traffic-related risks, and enhance freight network resilience, paving the way for advanced freight optimization strategies and sustainable urban development.

A Dynamic Road Damage Prediction Framework Using XGBoost and ArcGIS: Enhancing Flood Disaster Management in New Jersey

Haoxiang Liu, Dan Liu

Department of Management, Kean University, Union, New Jersey, USA

Abstract

Flood disasters significantly disrupt road infrastructure, posing challenges to transportation networks and disaster management systems. Traditional approaches to assessing flood-induced road damage often rely on static methods or heuristic models, which fail to adapt to the dynamic and uncertain nature of flood events. These methods are also limited in their ability to integrate real-time data and provide accurate, actionable insights for decision-making. To address these limitations, this study introduces a **Dynamic Road Damage Prediction Framework (DRDPF)** that integrates XGBoost and ArcGIS for flood disaster management, **focusing on dynamic risk assessment and geospatial visualization**.

The framework leverages XGBoost to predict road damage probabilities by combining static road and geographical features with dynamic weather and water level data derived from sliding time windows. Real-time IoT sensor data, including rainfall and water levels, is incorporated to ensure the model remains responsive to evolving flood conditions. ArcGIS is employed for spatial analysis and visualization, enabling the creation of dynamic heatmaps to identify high-risk road segments and prioritize recovery efforts effectively. This integration provides a data-driven and context-sensitive approach to flood risk management, ensuring timely and informed decision-making.

To evaluate the effectiveness of DRDPF, the study focuses on a flood-prone region in New Jersey, utilizing comprehensive data from historical and real-time sources. The experimentation framework includes spatial and temporal cross-validation to ensure model robustness and generalizability. The results highlight the framework's ability to accurately predict road damage probabilities and dynamically visualize risk distributions, demonstrating its potential to outperform traditional static methods in terms of adaptability and precision. By integrating XGBoost for predictive modeling and ArcGIS for geospatial analysis, the DRDPF offers a streamlined and scalable solution for flood disaster management, with significant implications for infrastructure resilience and emergency response planning.

Artificial Intelligence Policy in Lethal Autonomous Weapons Systems (LAWS)
across the United States, China, and the European Union

Janel Baxter, Dr. Anshu Arora, Dr. Amit Arora, Dr. Pradeep Behera

University of the District of Columbia, Washington, District of Columbia, USA

Abstract

This research explores the artificial intelligence (AI) policy and regulatory landscape surrounding Lethal Autonomous Weapons Systems (LAWS) in the United States, China, and the European Union, emphasizing the critical challenges of explainability and traceability. We aim to uncover how these influential world regions navigate the complex interplay of regulations, ethical problems, and strategic considerations that come with developing and deploying LAWS. By comparing their policies, we reveal the complex challenges of establishing international norms for autonomous weapons. This study highlights the ongoing tensions between leveraging technological advancements for military superiority and ensuring ethical oversight over lethal force. We integrate recent developments in AI governance and international security to provide a comprehensive view of the evolving landscape of LAWS regulation. Moreover, we discuss how differing national policies could shape global security dynamics and influence the future of warfare. By weaving together insights from technical, legal, and ethical perspectives, we hope to foster a more engaging conversation about responsibly governing AI-driven autonomous weapons systems in our increasingly interconnected world. Ultimately, this research is not just an academic inquiry; it's a call for collaboration among policymakers, ethicists, and technologists to ensure a safer future in military technology.

SOCIAL ROBOTICS AND HUMAN-ROBOT INTERACTION IN HEALTHCARE

Nkechi Iweriebor, Dr. Anshu Arora, Dr. Amit Arora, Dr. Bushra Saeed

University of the District of Columbia, Washington, District of Columbia, USA

Abstract

Social robotics and human-robot interaction (HRI) impact healthcare, improving patient outcomes and experiences. Current research trends inform the development of innovative social robotics and enhance human-robot interaction quality in healthcare. Healthcare robotics research have delved into robot types, benefits, challenges, robot acceptance, and ethical considerations in healthcare. This research examines the current state and future prospects of social robots in healthcare settings, focusing on elderly care, pediatric care, rehabilitation, and emotional-cognitive needs of patients. This research includes artificial intelligence, machine learning, and natural language capabilities. We analyze various robotic systems, including intelligent medical assistants, life-assisted nursing robots, and companion robots, highlighting their roles in enhancing patient care, autonomy, and quality of life. Providers can deliver personalized, compassionate care with empathy by introducing and integrating social robotics and HRI in current healthcare and the future. The study emphasizes the importance of safe, effective and ethical HRI, provides insights into the potential of social robotics and HRI to revolutionize healthcare delivery, focuses on managerial implications for healthcare institutions considering the integration of AI and robotic systems, formulates policy implications for regulating and standardizing the use of social robots in medical environments, and aims to guide future research and shape policies to ensure the ethical and effective implementation of social robotics in healthcare.

Emotional Artificial Intelligence in Customer Service: Can AI Truly Replace Human Empathy?

Tamaya Edwards, Dr. Anshu Arora, Dr. Amit Arora

University of the District of Columbia, Washington, District of Columbia, USA

Abstract

As the emerging technology of Artificial Intelligence (AI) continues to surpass expectations, taking over jobs by mechanizing tasks and substituting roles that would traditionally be performed by humans, the question as to how far AI can go to replace human capabilities arises daily. While we are witnessing AI successfully replicating the physical ability of humans in various fields, the real question lies in the possibility of it emulating the emotional aspect of humans. This study is based on an analysis of existing literature, collected from prior research on the ability of AI to replicate the emotions of humans. The synthesis of these findings, which include diverse research on Emotional Artificial Intelligence (EAI) with social robots, and how the EAI-based social robotics research can be integrated in customer service, aims to find if AI can truly imitate human empathy with depth and authenticity required for meaningful human interaction. The study investigates the main challenges, constraints, and improvements in, aims to determine the extent to which AI can incorporate empathy, and focuses on significant findings regarding EAI in building genuine human connections in service-based interactions.

Enhancing Agricultural Efficiency and Safety: The Synergy of AI, Robotics, and Human-Centered Smart Farming

Nicholas Brown, Dr. Amit Arora, Dr. Anshu Arora

University of the District of Columbia, Washington, District of Columbia, USA

Abstract

The integration of Artificial Intelligence (AI) and robotics is transforming agriculture, ushering in a new era of smart farming that prioritizes efficiency, precision, and safety. This research examines the evolving role of AI-driven robotics in optimizing key agricultural processes, including crop cultivation, pest management, soil monitoring, and resource allocation. Recent breakthroughs in computer science, sensor technology, and automation have led to the development of intelligent agricultural robots capable of executing precise, autonomous tasks. AI-powered analytics now provide farmers with real-time insights, ranging from weather predictions to soil health diagnostics, leading to improved yields, reduced resource waste, and enhanced sustainability.

By synthesizing insights from academic research and industry applications, this study presents a comprehensive analysis of the current advancements and future potential of AI and robotics in agriculture. A core focus is placed on human-centered design, ensuring that these technologies enhance user experience, safety, and long-term viability. This research explores critical questions, such as: How can AI-driven decision-making enhance the precision and efficiency of agricultural tasks? How do AI and robotics mitigate labor risks and improve farm safety? What role do autonomous robots and real-time data analytics play in accelerating farming operations? The findings suggest that continued advancements in AI and robotics will be key to addressing the growing global demand for food while promoting sustainable and resource-efficient farming practices.

Keywords: Smart Farming, Artificial Intelligence (AI), Agricultural Robotics, Sustainable Agriculture

A Multi-Objective Optimization Framework for ESG-Constrained Portfolios using Advanced Diversification Techniques

Ignacio Pericet, Amit Arora, Anshu Arora

University of the District of Columbia, Washington, DC, USA

Abstract

The growing emphasis on sustainable investing has prompted the need for innovative approaches to integrate Environmental, Social, and Governance (ESG) considerations into the portfolio construction process. This research introduces a multi-objective optimization framework designed to extract alpha while balancing financial performance and societal impact. Utilizing Upright's forward-looking ESG impact scoring, the study advances beyond traditional ESG methodologies by incorporating science-based net impact data to evaluate environmental, societal, governance, and financial dimensions. The framework leverages Upright's data to filter the investment universe, identifying the company's net impact. The Hierarchical Risk Parity (HRP) algorithm is then employed to construct a robust portfolio structure by ensuring risk is distributed evenly across asset clusters, mitigating estimation errors in covariance matrices and enhancing diversification. Finally, the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) optimizes for additional objectives such as expected return, Conditional Value-at-Risk (CvaR), and alignment with ESG criteria, providing a framework that not only meets sustainable finance objectives but also uncovers new avenues for alpha generation. The findings demonstrate how investors can navigate the complex trade-offs between financial and societal goals to construct portfolios that achieve both positive impact and financial performance.

The determinants of AI usage in the workplace: a comparative study between job-market entrants and young professionals

Leslie Lopez, Krupali Soni, Kayla Grant, Aaron van Klyton, Md. Golam Kibria

Kean University, Union, NJ, USA

Abstract

Artificial intelligence (AI) is becoming a game-changer in the modern workplace, offering transformative potential for productivity, decision-making, and innovation while also raising concerns about job displacement and diminished autonomy. For firms to maximize the use of AI in the workplace, it is imperative to understand how job candidates and early-career professionals adopt and adapt to such emerging technologies. This study determines the antecedents of individual AI adoption/usage in the workplace and evaluates the differences in perception between job-market entrants and young professionals regarding the use of AI in businesses. This generational comparison is novel and highlights important differences in how each group perceives and experiences AI in their job roles. The findings highlight the need for tailored AI training and education to address generational concerns, enhance adaptability, and support smooth AI-driven workforce transitions through informed HR strategies. The research model is grounded in the Technology Readiness Index and Task Technology Fit theories and will be tested using survey data. The data will be analyzed using structural equation modeling and the findings will contribute to a deeper understanding of the mechanisms underlying AI adoption in the workplace and provide both theoretical and practical implications for effectively navigating AI-driven workforce transformations.

Navigating the Blue Ocean and Economy: Renewable Energy Solutions and Sustainable Maritime Practices

Josh Esteban, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Abstract

With the rising concerns over greenhouse emissions and climate change, this paper will examine the incorporation of renewable energy and electric propulsion technologies in maritime shipping by focusing on green alternatives to fossil fuels to combat those problems. Key solutions include wind-assisted propulsion such as rigid sails and kite sails, solar driven systems, hybrid/dual fuel systems, and low carbon fuels such as LNG, clean ammonia, biofuels, green hydrogen, and methanol. However, these solutions do not come without challenges and barriers. This paper will go over barriers such as energy storage, costs of utilizing sustainable energy, infrastructure gaps, and challenges in regulation. The study will also highlight how government aid through subsidies, tax incentives, and grants will play a crucial role in adopting these renewable and sustainable solutions. By going over research on public and private partnerships and pilot programs, this study will demonstrate how investments can reduce costs and drive innovation and help lower those barriers. By going over the pros and cons of these renewable solutions by analyzing their viability, economic effects, impact on the environment, and scalability, this study will provide valuable insights for advancing sustainable shipping, meeting regulations, and global decarbonization within shipping.

Urban Planning: Implementing Accessibility in City Design

Eman Hermas, EunSu Lee

New Jersey City University, Jersey City, NJ, USA

Abstract

This research examines the benefits of accessibility in urban areas, specifically through the means of walkability and reliable public transportation. Areas that have walkable cities are found to have positively impacted its residents' mental and physical health. However, not all walkable cities are made equal. The availability and quality of public transportation are modes of access that are integral in a community. Many residents also may also opt to utilize bikes to get to their destination, which can't be done in certain cities that aren't designed for bike lanes. It is important for urban planners to consider other modes of transportation when designing cities, not just cars.

Efficacy of Phishing Remediation: An Undergraduate Honors Thesis

Chandra Dhimal, Michael Douglas

Millersville University, Millersville, PA, USA

Abstract

This study examines phishing susceptibility among students and staff at Millersville University. In Spring 2025, 600 students and 100 staff members will be targeted by a simulated phishing attack using the same email from Jamie Thorpe's 2017 study. Participants will complete a survey afterward to assess their responses and understanding of phishing tactics. The goal is to improve cybersecurity awareness and develop more effective educational strategies to prevent phishing in university settings.

The research applies the Social Engineering Life Cycle framework, which identifies stages of social engineering attacks: information gathering, engagement, exploitation, and execution. By examining these stages, the study aims to determine which are most influential in shaping participants' behaviors and how psychological factors such as trust, urgency, liking, and authority impact decision-making. These insights will help create tailored interventions to address specific vulnerabilities within the university community.

Building on Thorpe's 2017 work, which analyzed phishing emails using machine learning techniques, this study focuses on human behavior. Thorpe's research examined the technical aspects of phishing emails, while this study explores how individuals respond to these emails. Integrating psychological factors, the research adds a new layer to understanding phishing susceptibility.

Additionally, the study introduces targeted educational interventions. Based on participant responses, customized training programs will be developed to address identified weaknesses, such as overreliance on authority or susceptibility to urgency. This approach aims to improve cybersecurity awareness and resilience in academic environments.

By comparing the results with Thorpe's findings, this research provides a longitudinal perspective on phishing response rates and user behavior. The study aims to bridge the gap between technical analysis and human factors, ultimately enhancing cybersecurity in university communities and contributing to the broader field of cybersecurity education.

Medical Malpractice: Legal, Insurance, and Ethical Implications. Inspired by a True Story.

Claire Coons

Millersville University, Millersville, PA, USA

Abstract

Medical malpractice, sometimes referred to as medical negligence, occurs when a healthcare provider deviates from the accepted standards of treatment for a patient, leading to injury or sometimes death. Often, this form of negligence will cause injuries that a patient will deal with for the rest of their life. Severe cases of medical malpractice can lead to death. Medical malpractice drastically impacts a patient's life, including legal, insurance, and ethical implications. This paper was inspired by a true story that I experienced in the latter of 2024. My research includes a brief history, authority of the individual states, the legal elements, insurance coverage and tax considerations, types of insurance policies, ethical concerns, and a case study.

Rich Picture Analysis for Health Policy

Akira Dasgupta¹, Kuan-Lung Chen²

¹Riverside High School, Leesburg, VA, USA. ²The George Washington University, Washington, DC, USA

Abstract

Rich Picture Analysis is a form of soft systems methodology. It provides a way to move from a state of meta-confusion—characterized by profound uncertainty and lack of clarity—to a clearer understanding of key themes. Participants use rich pictures as a sketching method to represent and visualize a situation.

This research offers an overview of the rich picture technique and applies it specifically to health policy, which is part of the broader field of public health. According to the CDC Foundation, public health is dedicated to protecting and promoting the health of populations. This involves preventing disease, encouraging healthy lifestyles, and responding to health threats through research, policy development, and community interventions. The goal is to improve the overall well-being of communities rather than focusing solely on individual patients.

In this study, we utilize rich picture analysis to explore the various factors that influence the impact of health policy on a community.

Euthanasia Rates Among Companion Animals in Los Angeles County Compared to Income Level

Isabelle Darcy

Bridgewater State University, Bridgewater, MA, USA

Abstract

Euthanasia rates of companion animals across the United States are simply becoming an unacceptable reality. Animal control centers and shelters are continually adding to the statistics of animal euthanasia. In 2024, Los Angeles County euthanized over four thousand companion animals. A GIS map will be created to analyze the Los Angeles Animal Care and Control data and household income provided by the United States Census. This research will offer insights into whether income is correlated to higher euthanasia rates. This study will identify areas needing extra support to reduce euthanasia, thereby contributing to the goal of implementing better practices.

Evaluating Student Experiences with the Class Waitlist System at NDSU
Samantha Hendrickson, Anna Kessel, Kaylyn Haug, Rylee Stotz, Alison Hinze
North Dakota State University, Fargo, ND, USA

Abstract

This research explores the perceptions, challenges, and opportunities for improvement in the class waitlist system at North Dakota State University (NDSU). Using a mixed-methods approach, we surveyed 41 students within the College of Business to examine their experiences with the Campus Connection waitlist system. Key areas of focus include awareness of the system, communication effectiveness, emotional impacts, and the role of external influences, such as class formats and professor preferences. Findings reveal significant dissatisfaction with communication and a strong preference for real-time updates, such as position tracking. Recommendations for system improvement emphasize transparency, enhanced notifications, and user-friendly features to support students' academic planning. This study provides actionable insights for institutional decision-makers to enhance the waitlist process, ultimately reducing stress and improving student outcomes.

The Convergence of Generative AI and Quality Assurance: Enhancing Pharmaceutical Supply Chain Resilience.

Aoibhin Lavery, Elif Kongar

University of New Haven, West Haven, CT, USA

Abstract

Building more resilient supply chains can deliver consistent, high-quality pharmaceutical products, even in the face of unforeseen disruptions or fluctuating market demands. Today's emerging AI-driven solutions can significantly boost these efforts. For instance, integrating AI into pharmaceutical operations has been effective in enhancing predictive maintenance in manufacturing. This minimizes downtime, ensures compliance with safety and quality regulations, and keeps equipment and processes running smoothly. Beyond manufacturing, AI-powered automation streamlines documentation processes, while predictive analytics identifies potential risks, such as shortages or quality issues, before they escalate, enabling timely and effective interventions. With this motivation, this research investigates the impact of generative AI on quality assurance and regulatory compliance within pharmaceutical supply chains. Specifically, we explore the integration of **Analytical AI (AAI)**, **Autonomous AI (AUI)**, **Causal AI (CAI)**, **Deep Learning (DL)**, **Generative AI (GAI)**, and Machine Learning (ML) to address critical challenges in these areas. To structure our investigation, we developed an Ishikawa diagram to systematically categorize the causes of non-compliance in pharmaceutical supply chains.

Through an **analysis of FDA inspection citations with over 250 thousand citations**, an extensive literature review, and insights from industry professionals, we identified four primary categories, viz., *Compliance*, *Documentation*, *Operations*, and *Quality*, that collectively account for over 80% of pharmaceutical compliance issues. Our findings also highlight that *while each AI tool has its unique role in enhancing supply chain resilience, ML emerges as the most impactful in addressing critical challenges in quality assurance and regulatory compliance*. AAI and DL follow closely behind in their contributions, while CAI appears to be the least utilized. Further research is needed to explore CAI's potential applications and benefits in this context.

This study underscores the transformative potential of AI in strengthening pharmaceutical supply chains, paving the way for enhanced quality, efficiency, and resilience in the face of complex challenges.

Resilient Supply Chains, Resilient Communities: An AI-Powered Emergency Response Framework

Aarya Kshetri, Nahian Karim, Sneha Anumandla, Elif Kongar

University of New Haven, West Haven, Connecticut, USA

Abstract

AI technologies such as Machine Learning (ML) and Predictive Analytics (PA) are transformative tools for managing supply chain risks, enhancing resilience, and addressing vulnerabilities. These technologies hold great potential not only for predicting disasters during early preparedness but also for facilitating effective response and recovery efforts. For humanitarian organizations, leveraging such capabilities is crucial to allocate resources where they are needed most, while addressing challenges such as communication breakdowns, coordination gaps, and delays in planning and aid delivery.

Local non-profit organizations are particularly well-suited to respond swiftly and effectively in disaster scenarios due to their proximity to affected communities and the trust they have established. However, these organizations often face significant barriers, including limited financial resources and insufficient access to advanced technologies, which can hinder their ability to make timely, informed decisions during crises.

To address these challenges, this study introduces **SHIELD (Supply Hubs Integrated Emergency Logistics Delivery)**, an open-access data-driven disaster planning and management tool designed to enhance humanitarian relief supply chains. SHIELD is a dynamic dashboard, built using R and Python, that leverages machine learning and predictive analytics to integrate non-profit organizations with suppliers of essential goods and services. Using Tableau for visualization, the platform aims at consolidating real-time data, generates predictive models, and delivers actionable insights to support effective decision-making. SHIELD's functionality is demonstrated using a data-driven example, showcasing its ability to strengthen disaster preparedness and recovery efforts.

This approach minimizes disruptions, empowering non-profit organizations to provide timely and effective relief to those in need. By combining AI technologies with data-driven analytics, SHIELD underscores the transformative role of AI in creating sustainable, equitable, and impactful disaster management solutions.

Generative AI in Omni-Channel Marketing

Aarya Upadhyay, Ellen Belitzky

University of New Haven, West Haven, CT, USA

Abstract

Generative AI is at the forefront of transforming omnichannel marketing, focusing on its ability to elevate customer experiences across different platforms. By leveraging insights from McKinsey research and a Nike case study, this research identifies key mechanisms through which generative AI drives hyper-personalization and streamlines customer engagement. The analysis highlights known benefits such as efficiency gains and predictive analytics. It also addresses challenges such as ethical considerations, data privacy concerns, and the complexities of integrating across multiple platforms.

To address generative AI as a catalyst for omni channel marketing transformation, this research takes a dual approach: first, evaluating the impact of AI-generated content on customer satisfaction and loyalty; second, identifying frameworks that help businesses overcome adoption barriers. By examining the influence of AI-driven personalization, predictive analytics and real-time customer engagement provides a foundation for integrating generative AI into marketing strategies effectively.

Findings are expected to highlight how generative AI can create seamless, value-driven customer journeys while emphasizing the need for robust ethical guidelines and adaptable operational strategies. By balancing the benefits and challenges, this research provides suggestions to effectively implement generative AI in omnichannel marketing, driving innovation, enhancing customer experiences, and fostering sustainable business growth.

AI for Cost Reduction in B2C E-Commerce Supply Chains

Simran Harjani, Ellen Belitzky

University of New Haven, West Haven, CT, USA

Abstract

Artificial Intelligence (AI) tools empower e-commerce businesses with advanced capabilities such as demand forecasting, real-time supply chain visibility, and predictive analytics, enabling more efficient decision-making and significant cost reductions. Business to consumer (B2C) e-commerce focuses on direct transactions through online platforms. Ineffective use of AI can leave businesses behind, missing opportunities and facing inefficiencies. Mastering predictive analysis is essential, not optional, to thrive in today's competitive market.

This research explores how artificial intelligence is revolutionizing the supply chain by addressing limitations of traditional e-commerce methods to meet consumer demand. The e-commerce market is projected to reach approximately US\$4 billion by 2025. The integration of AI into supply chain management presents a crucial opportunity for businesses to optimize operations and achieve sustainable growth given this exponentially increasing market.

AI tools, such as Blue Yonder, IBM Supply Chain, and Llamasoft Supply Chain Guru, may be utilized for supply chain cost reduction and internal operational efficiency through different algorithms and metrics. Nearly 49% of technological leaders have integrated AI in their core business strategy in 2024. With effective implementation, significant growth in sales and reduced costs of logistic operations could be realized in 2025. The research will demonstrate how these advancements not only improve overall competitiveness, but also pave the way for a future in e-commerce where AI services business efficiency and innovation to solve growing market challenges.

Pitfalls of Over-Dependency on AI in Financial Markets

Nikhita Rane, Ellen Belitzky

University of New Haven, West Haven, CT, USA

Abstract

As artificial intelligence (AI) grows more prevalent in financial markets, concerns regarding over-reliance on AI-driven systems are rising. This research investigates the possible risks of over-reliance on AI, particularly during situations of market stress. One significant issue is AI-induced stress selling, in which algorithmic systems built to respond to market signals aggravate panic-driven sell-offs, hence increasing market volatility and liquidity problems. Furthermore, algorithmic decision fatigue might arise from AI systems' incapacity to account for the complex human emotions and illogical behaviors that frequently accompany market stress, resulting in inefficient decision-making.

The interaction of AI algorithms across different platforms raises concerns about market feedback loops, in which automated systems respond to one another in ways that amplify market fluctuations, potentially contributing to flash crashes or surprise market collapses. This study underlines the importance of striking a balance between AI automation and human monitoring, especially during uncertain times, to avoid AI-driven market disruptions. Furthermore, it explores the ethical and regulatory concerns required to prevent over-reliance on AI, guaranteeing that human intervention can reduce the dangers of systemic breakdowns in an AI-driven financial system.

Emotional Intelligence as a Measure of Potential Academic Success Angelina Zuger, Kubra Banu, Joash Kawal, and Kabhilan Shanmugasundaram Business Analytics Master's Program, Clark University School of Business, Worcester, MA 01610

Angelina Zuger, Kubra Banu, Joash Kawal, Kabhi Shanmugasundaram

Clark University, Worcester, MA, USA

Abstract

Measures of academic performance, often used as standards of admission, include measures of cognitive ability, such as IQ, grade point average (GPA), SATs, and GREs, are limited in assessing a student's ability to navigate a rigorous academic setting. Recent studies show that these measures are not accurate when evaluating academic performance, since such scores provide a limited view of a student, excluding the behavioral and emotional aspects of their personality. In the last ten years, Emotional intelligence (EI) has become known as a potent predictor of both academic and career achievement. EI focuses on personality characteristics that involve coping strategies, socio-emotional skills, and stress management, attributes essential to navigating academically competitive graduate programs. Can measures of Emotional Intelligence be accurate predictors of academic success in graduate school, by comparing measures of student success (standardized test scores) to measures of Emotional Intelligence? To address this hypothesis, Linear Regression analysis, paired with a Branched Decision Tree Model, were used to analyze a dataset compiled from a paper by MacCann et al., 2020 of the University of New South Wales, titled "Emotional Intelligence predicts Academic Performance." The results of the Regression Analysis demonstrate that none of the models that were run fit well within the given dataset. Despite large and significant F statistics in the models, the adjusted R-squared values show that there is little explanatory power for the variance in the data. The results of the Branched Decision Tree model showed that variance (vi) and population (n) were the most important variables in the construction of the tree, with importance values of 5.73 and 4.36, respectively. The conclusions of this study demonstrated a very weak association between EI and academic success, partly due to the limitations of the dataset, notably, the high variance and a differing population sizes within the cohort.

PROBLEMS FACED BY INDIAN EXPORTERS TO EXPORT PSYLLIUM HUSKS TO THE USA

Dharmi Shah, EunSu Lee

New Jersey City University, Jersey City, NJ, US

Abstract

Psyllium, or ispaghula, is the common name used for several members of the plant genus *Plantago* whose seeds are used commercially to produce mucilage. Psyllium is mainly used as a dietary fiber to relieve symptoms of both constipation and mild diarrhea, and occasionally as a food thickener. Psyllium is a soluble fiber used primarily as a gentle bulk-forming laxative in products such as Metamucil. It comes from a shrub-like herb called *Plantago ovata* that grows worldwide but is most common in India. Each plant can produce up to 15,000 tiny, gel-coated seeds, from which psyllium husk is derived. The soluble fiber found in psyllium husks can help lower cholesterol. Psyllium can help relieve both constipation and diarrhea, and is used to treat irritable bowel syndrome, hemorrhoids, and other intestinal problems. Psyllium has also been used to help regulate blood sugar levels in people with diabetes. When psyllium husk meets water, it swells and forms a gelatin-like-mass that helps transport waste through the intestinal tract. Several large population-based studies also suggest that increased fiber intake may reduce the risk of colon cancer, but other studies have been conflicting.

LEADERSHIP AND TEAM SPORTS

Denish Patel¹, EunSu Lee²

¹Dallas Baptist University, Dallas, TX, USA. ²New Jersey City University, Jersey City, NJ, USA

Abstract

This research paper thoroughly explores how leadership is crucial in team sports. It explains that leadership can guide actions, drive team success, and foster teamwork. The paper looks at different leadership styles, how they affect a team's performance, and what qualities make a good sports leader. It compares leadership to the foundation of a building, emphasizing its importance. The research aims to provide valuable insights for coaches, athletes, and sports fans. By breaking down these complex ideas, it shows that leadership plays a key role in achieving success, teamwork, and accomplishments in team sports.

If they Can, Why Can't We? Comparing the United States' and European Union's Approaches to Artificial Intelligence Governance and Regulation

Muano Nevhudzholi, Mariam Sanogo, Dr. Anshu Arora, Anshu Arora

Abstract

This research explores the different approaches to artificial intelligence (AI) governance and regulation in the United States (US) and the European Union (EU). The EU has taken a proactive standpoint with the comprehensive AI Act. In contrast, the US has adopted a more decentralized, sector-specific approach to AI regulation. Although there is no unified “US AI Act or Policy,” the US has been shaping its AI governance framework through executive actions aimed at state-level legislation to address specific AI-related concerns. Congress continues to work on measures to regulate AI at a federal level, and these initiatives reflect a varied approach to managing the complexities of AI governance and regulations. We analyze the implications of these divergent strategies for ethical AI development and international competitiveness by exploring how the EU's regulations may influence global norms, potentially impacting US companies operating in the European markets.

Air Quality Improvements: A Catalyst for Better Respiratory Health

Sara Rehman, Anvita Patne

Clark University, Worcester, Massachusetts, USA

Abstract

The relationship between air quality and public health is one of the most critical areas of environmental research. This paper examines publicly available data to assess whether there is a correlation between air quality improvement initiatives funded by the United States Environmental Protection Agency and changes in respiratory health outcomes within affected communities. The current study assesses the effectiveness of such initiatives based on combined data from the EPA air quality indices and the CDC respiratory health statistics.

Our approach uses straightforward statistical methods to analyze trends in air pollutant levels, such as particulate matter (PM_{2.5}) and ground-level ozone, alongside rates of respiratory conditions, including asthma and chronic obstructive pulmonary disease (COPD). By examining publicly available data from communities targeted by EPA-funded programs, we aim to identify patterns and assess the direct impacts of these initiatives.

Preliminary data show that areas with EPA-funded air quality programs provide measurable air quality improvement and a reduction in associated health outcomes in the form of respiratory-related hospitalization and emergency room visits. This constitutes a good argument for understanding the environmental health burdens and taking some interventions through readily available data.

This study highlights the role of air quality improvement programs in ensuring the protection of public health, especially within more at-risk communities. It also emphasizes the potential of publicly available datasets to be used for policy decisions and encourage transparency in environmental health research.

Our work hopes to provide actionable insights for policymakers to conceive and implement sustainable and equitable air quality initiatives that improve community health outcomes. Since the focus is on publicly available data, this can hopefully be a replicable framework for further studies in this intersection of environmental policy and public health.

Author Index

Abdou, Khaled, 474
Abdul, Chaza, 296
Adams, Tom, 544
Adams, Brian, 478
Adams-Harmon, Dawn, 524
Addeo, Eric, 567
Addo, Sampson, 503
Adobor, Henry, 542
Agarwal, Anu, 572
Aggarwal, Anil, 514
Alatorre, David, 583
Altay, Nezih, 548
Amendah, Eklou, 531
Anoruo, Emmanuel, 469
Antonucci, Yvonne Lederer, 553
Anumandla, Sneha, 601
Arhin, Kofi, 110
Arinze, Bay, 90
Arize, Augustine, 473
Arora, Amit, 494, 587, 588, 589, 590, 591, 608
Arora, Anshu, 494, 587, 588, 589, 590, 591, 609
Asonibare, Toyin, 163
Awasthi, Anjali, 566
Ayodele, Olanrewaju, 338
Badve, Ameya, 527
Bahalkeh, Esmaeil, 522
Balakrishnan, Anantaram, 536
Balan, Shilpa, 487
Balasubramanian, Ganesh, 561
Banu, Kubra, 605

Bates, Diane, 507
Baxter, Janel, 587
Behera, Pradeep, 587
Belitzky, Ellen, 525, 602, 603, 604
Bezawada, Ramakrishnainh, 527
Bhandare, Kajal, 487
Bhatt, Ganesh, 163, 402
Bolinsky, Olivia, 468
Bougioukos, Vasileios, 571
Branch, Benjamin, 495
Brown, Richard, 540, 544
Brown, Nicholas, 590
Bunyaratavej, Kraiwinee, 488
Burns, Ian, 579
Carbo, Jerry, 570
Cariati, Francesco, 571
Casey, Haleigh, 14
Casson, Micheal, 495
Chang, Aichih, 568
Chauhan, Satyaveer Singh, 535
Chen, Ja-shen, 360
Chen, Jing, 559
Chen, Bintong, 559
Chen, Kuan-Lung, 597
Chiang, Bea, 507
Cho, Youngwoo, 581
Choi, Heeick, 467
Chou, Shih, 510
Chowdhury, Tilottama, 529
Chuang, Michael, 489
Chukwuogor, Chiaku, 469
Coffta, Michael, 517

Coons, Claire, 596
Dalmadge, Cretson, 510
Darcy, Isabelle, 598
Dasgupta, Akira, 597
Dasgupta, Subhasish, 489
Deal, Zachary, 580
Desnoyers, Neil, 508
Dhanapal, Kaverinathan, 538
Dhimal, Chandra, 595
Doh, Jonathan, 543
Douglas, Michael, 14, 518, 595
Duman, Gazi, 485
Ebrahimpour, Maling, 519
Edwards, Tamaya, 589
Eevani, Sri Ramesh, 497
Ejiaku, Samuel, 402
El-Rayes, Nesreen, 568
Esteban, Josh, 593
Everett, Conner, 578
Fang, Jianing, 479
Fang, Jin, 486
Fang, Liping, 552
Farr, Lori, 71
Ferrerias-Perez, Jennifer, 100
Ferris, Kathleen, 506
Ferry, Nevin, 577
Fischer, Dov, 477, 482
Flasher, Renee, 468
Fortino, Andres, 139, 193, 208
Fowler, Danielle, 514
Gabaree, Simon, 574
Galdamez, Estefany, 528

Galkina, Elena, 565
Gallagher, Grace, 29
Gefen, David, 498, 546
Givens, Lindsay, 556
Gopalan, Ram, 307
Gorkhali, Anjee, 491
Grant, Kayla, 592
Gray, Damon, 273
Green, Jarvis, 495
Gunasekaran, Angappa, 564
Guo, Ge, 447
Gupta, Surendra, 485
Gupta, Parvinder, 541
Gupta, Shubham, 562, 563
Gupta, Ajay, 572
Haghighi, Aida, 493
Hahn, Eugene, 543
Hales, Douglas, 496
Harithsa, Jyothsna, 110
Harjani, Simran, 603
Harper, Alisha, 492, 573
Hasan, Rashedul, 484, 504
Haug, Kaylyn, 599
Helman, Deborah, 567
Hendrickson, Samantha, 599
Hermas, Eman, 594
Higgs, Bryan, 503, 505, 509
Hikmet, Neset, 495, 521
Hinze, Alison, 599
Hong, Jae-Dong, 53, 432
Huang, Sawyer, 139
Issa, Samar, 471

Issa, Hussein, 471
Iweriebor, Nkechi, 588
Jackson, Jonathan, 512
Jacobsen, Amal, 131
Jaska, Pat, 549
Jin, Yan, 555
Jones, Patricia, 71
Joo, Sunghoon, 467
Kadagadakai, Ankith Anil, 576
Kalal, Shreyash, 238
Kapoor, Latika, 487
Karam, Ali Gamaleldin, 571
Karim, Nahian, 601
Karimi, Haleh, 492
Kashyap, Rajiv, 252
Kawal, Joash, 605
Kesavaprakash, Vignesh, 562
Kessel, Anna, 599
Khan, Zarak, 252
Khedr, Mohamed Elshahhat Ali, 571
Kibria, Md. Golam, 490, 592
Kim, Sung Ik, 475
Kim, Seung-Lae, 530, 539
Kimerling, Lionel, 572
Kirac, Emre, 502
Kline, William, 540, 544
Klyton, Aaron van, 592
Kongar, Elif, 485, 600, 601
Kouhizadeh, Mahtab, 519
Kowalski, Kellyann, 547
Kshetri, Aarya, 601
Kumar, Archana, 531

Kumar, Sanuj, 538
Kyper, Eric, 14
Lahmiri, Salim, 535
Langella, Ian, 570
Lash, Michael, 500
Lavery, Aoibhin, 600
Lee, Gyeong Jun, 493
Lee, Howard, 472
Lee, Seung Won, 467
Lee, EunSu, 576, 581, 582, 593, 594, 606, 607
Lennon, Mark, 570
Lewis, Aleisha, 524
Li, Tao, 466
Li, Gang, 536
Li, Yang, 515
Li, Kyle, 500
Liao, Tung-Shan, 569
Lin Lung, Shan-Hai, 569
Litsiou, Konstantia, 571
Liu, Dan (JMU), 537
Liu, Dan (Kean), 558, 584, 585, 586
Liu, Yiyang, 585
Liu, Haoxiang, 586
Liu, Huiling, 532
Liu, Xiangrong, 572
Lopez, Leslie, 592
Lowrance, Roy, 139, 208
Mahyari, Ehsan, 534
Mai, Khiem Dinh, 360
Malaga, Ross, 513
Malindretos, John, 473
Marchant, Alli, 1

Mariani, Ronda, 296, 334, 517
Mascarenhas, Briance, 523
Mehta, Khushi, 419
Meklachi, Taoufik, 478
Menash, Patrick, 495
Menk, K Bryan, 1
Meyer, Christopher, 549
Milligan, Kimberly, 495
Misra, Ram, 515
Mitra, Prakarsha, 64
Moatari-Kazerouni, Afrooz, 493, 553, 554
Mohamed Ismail, Mohamed Wahab, 552
Motiwalla, Luvai, 501
Murshed, Feisal, 529
Muszynski III, Richard, 476
Mutha, Pavan, 550
Nafar, Reza, 535
Nag, Rajiv, 497
Nagarkar, Pradnya, 572
Nan, Wenxiu, 466
Narayanaswamy, Subramanian, 470
Ndu, Ikechukwu, 469, 481
Nevhudzholi, Muano, 608
Ogliari, Emanuele Giovanni Carlo, 571
Olsson, Amelia, 349
Ono, Takahiro, 572
Orenstein, Penina, 565
Ozbay, Gulnihal, 495
Ozkul, Ahmet, 520
Oztekin, Asil, 484, 501, 504
Pai, Dinesh R., 522
Pakdil, Fatma, 519

Pal, Raktim, 548
Partovi, Fariborz, 486
Patel, Denish, 607
Patel, Kushal, 238, 252
Patne, Anvitaa, 609
Patrichuk, Nicholas, 41
Pearson, Kim, 507
Pelaez, Alexander, 41, 131, 238
Penn Mekile, Richelyn, 546
Pericet, Ignacio, 591
Petsas, Iordanis, 483
Pinter, Janos, 533
Powell, Loreen, 296
Powell, Gwendolyn, 296
Prakash, Chandra, 550
Pulimoon, Monisha, 507
Puranam, Kartikeya, 523
Puranam, Srinivasa, 526
Qosej, Triumf, 498
Rafique, Raza, 551
Rane, Nikhita, 604
Raven, AJ, 233
Rehman, Sara, 609
Reyes, Joselyn, 582
Reyes, Pedro, 549
Rhee, Shanggeun, 545
Rossborough, Isabella, 574
Roy, Abhishek, 561, 563
Saeed, Bushra, 588
Saghezchi, Amirhossein, 530
Saini, Sajan, 572
Saleh, Feras, 566

Salmani, Yasamin, 486
Samosky, Serafina, 1
Samuel, Jim, 131, 238, 252
Sanogo, Mariam, 608
Sarkar, Piyal, 552, 563
Schaer, Oliver, 539
Schwartz, Jessica, 71
Schwartz, Tamara, 577, 578, 579, 580
Sebastianelli, Rose, 483
Sengupta, Kaushik, 131
Serna-Otalvaro, Samuel, 572
Shah, Dharmi, 606
Shah, Samir, 90
Shaikh, Musa, 550
Shang, Guangzhi, 537
Shanmugasundaram, Kabhi, 605
Shen, Philip, 100
Sherwani, Sahar Khan, 252
Shi, Jim, 568
Shi, Zhengzhong, 301
Shim, Sung, 516
Shrestha, Asim, 491
Sizemore, Caitlin, 573
Slavin, Nathan, 479
Soltaninejad, Mahsa, 539
Song, Hakjoon, 467
Soni, Krupali, 592
Sridhar, Karthik, 527
Stagliano, A. J., 480
Stamatopoulos, Theodoros, 473
Stein, Eric, 554
Stoner, Claire, 1

Stotz, Rylee, 599
Sullivan, Nancy, 525
Swanson, Jennifer, 547
Talaiei, Nolan, 501
Tamimi, Nabil, 483
Tan, Glynnis, 545
Tarcn, Menderes, 521
Tate, Chris, 495
Thomas, Ellen, 511
Tolbert, Devon, 575
Unger, Beth, 572
Upadhyay, Aarya, 602
Upatising, Benjavan, 519
Valian, Haleh, 499
Vidalis, Sofia, 338
Visich, John, 549
Walters, David, 567
Wang, Shengbin, 557
Wang, Tianyu, 208
Wang, Jinchang, 23, 124
Wang, Min, 560
Wolfe, Wojtek, 526
Xue, Yuan, 489
Yih, Yuehwern, 519
Yung, Jung Ah, 524
Zejnnullahi, Rreze, 513
Zhang, Drew, 519
Zhang, Cong, 514
Zhang, Chuanqian, 473
Zhao, Shaoping, 349
Zhou, Fuqin, 568
Zhu, Yuanyumeng, 584

Zhu, Hongwei, 501

Zhu, Xiaorui, 500

Zuger, Angelina, 605