ICEC · KITS · KIISS

ICEC 2026

The 27th International Conference on Electronic Commerce

Charting the Future of AX Ecosystems and AI-Driven E-Commerce

Date: **May 27**th-**29**th, 2026

Venue: Science and Technology Convention Center, Seoul, Korea

Hosted by International Center for Electronic Commerce, Korea Society of IT Service, Korea Intelligent Information Systems Society

Co-Hosted by Institute of Information & Communications Technology Planning & Evaluation

Sponsored by AiMX, KT, LG CNS, NICE Information Service, STA Consulting, VESSL AI, NAVER, NongHyup Bank, DaiShin Information & Communication, DOUZONE BIZON, Digital Content Group, LOGPRESSO, Lotte Members, Metanet Digital, Belltechsoft, Samsung SDS, Samsung Card, Seoul Metropolitan Government, Selim TSG, Korea Software Financial Cooperative, ITCEN ENTEC, MTDATA, YoungLimWon Soft Lab, OKESTRO, ALL4LAND, WAPLUS, Woori IT, WISEITECH, INFINIQ, National IT Industry Promotion Agency, KCA, Colosseum, Clean & Science, KIM Venturous, Tara TPS, Pasqal, Korea Institute for Advancement of Technology, Korea Association for Service, K water, Korea Fiscal Information Service, National Information Society Agency, Hansung University, KEPCO KDN, Hoseo University

ICEC International Center for Electronic Commerce

한국IT서비스학회 KOREA SOCIETY OF IT SERVICES

한국지능정보시스템학회 Korea Intelligent Information Systems Society

IITP 정보통신기획평가원

2026 –Wed 27th MAY PROGRAM

W1	**Live Streaming & Social Media Engagement / Session Chair : Dongyeon Kim (Hanyang Univ.)**	
13:20 ~ 14:40	To narrate or to list? The interplay of avatar type, presentation style and time-of-day effect in virtual livestreaming interaction	Xiaoxiao Gong and Jinhong Wang (Guizhou Univ.)
	Understanding Paid Subscription Behavior in YouTube Memberships: The Role of Emotional Attachment	Yen-Hao Hsieh, Ru-Yi Lai, Chih-Wei Hsu, Pei-Syuan Jiang, Pei-QiWu,Yi-ChingHsieh and Zheng-ZheChen (National Taichung Univ. of Science and Technology)
	Exploring the Thematic Structure of Live Streaming Commerce: A Bibliometric Analysis using LDA	Yu-Shan Huang and Chen-Hao Huang (National Cheng Kung Univ.)
	AI-Inferred Personality Traits and Humanlikeness in SoftwareReviews: Evidence from Collaboration Tool Platforms.	Sunnyoung Lee (Dongguk Univ.), Hsiu-Ping Lin and Jongwook Woo (California State Univ. Los Angeles)
	When Interaction Matters: Learning Effect of Knowledge-Sharing Live Streaming	Hanqiong Wu, Zhengzhi Guan (Beijing Normal-Hong Kong Baptist Univ.) and Bosen Lin (University of Nottingham Ningbo China)
W2	**Advanced Recommender Systems & Marketing / Session Chair : Sunghan Ryu (Univ. of Nottingham, Ningbo)**	
14:50 ~ 16:10	A Text Mining-based Analysis of Consumers' Online Pharmaceutical Purchasing Preferences	J. L. Xue and J. X. Yang (Anhui Univ. of Science and Technology)
	When AI Replies to Complaints: How Differences in Large Language Models Shape Consumers' Perceptions of Fairness and Brand Trust	Li Suo (Huizhou Univ.), Sijing Yang, Magnum Man-Lok Lam and Christina Wing-Yan Wong (The Hong Kong Polytechnic Univ.)
	Warmth over Realism: Why Affective Likeability Trumps Technical Anthropomorphism and Expertise in Virtual Live-Streaming	Tien-Dat Bui, Chih-Hsiung Hu (National Formosa Univ.), Tien-Dat Nguyen (National Formosa Univ. & Van Hien Univ.) and Chih-Chin Liang (National Formosa Univ.)
	From Expansive to Suppressive Escape: Emotional Exhaustion and User Disengagement on Immersive Digital Platforms	Peng He (Kookmin Univ.) and Jing Zhang (Kookmin Univ.)
	From Online to Offline Visit: Exploring Cognitive Fit Effect of AI-Streamer and Background	Weiyu Dai, Xusen Cheng, Zhangtong Gu, Yijun Yan and Lu Gao (Renmin Univ. of China)
	MORE: A Multi-Agent Approach for Multi-Objective E-commerce Recommendation	Yuxiang Pang and Di Wang (Xi'an Jiaotong Univ.)
	Learning When to Enrich: A Human-AI System for Dynamic Capability Development in Demand Forecasting	Bongsug Chae (Kansas State Univ.) and Eunhye Olivia Park (Gachon Univ.)

2026 –Wed 27th MAY PROGRAM

W3	**Generative AI & LLM Applications / Session Chair : Sung-Byung Yang (Kyung Hee Univ.)**	
16:20 ~ 17:40	Beyond the Classroom: A Task-Technology Fit Analysis of GenAI Speaking Tutors in Autonomous EFL Learning	Pei-Hung Ju, Tzu-Yu Lin, Zhi-Xuan Wu, Shang-Wen Huang, Ting-Chien Chiu(National Kaohsiung Univ. of Science and Technology), and Hsiao-Lan Wei (National Taiwan Univ. of Science and Technology)
	Dynamic Modality Weighting with Confidence-Based Fusion in Resource-Constrained Settings	Akshath H. P., Chaitra Reddy B., Mansa K. and Sunil Kumar V. (Nitte Meenakshi Institute of Technlogy)
	Examining the Developmental Trajectory of AI in Consumer Research: A Citation Network Analysis	Ching Yung Lin, Chen-Hao Huang and Tai-Yu Lin (National Cheng Kung Univ.)
	From Text to Presence: Mechanisms of Generative AI Comments in Social Media Engagement	Yijun Yan, Xusen Cheng and Changyu Zhou (Renmin Univ. of China)
	AI-Generated Review Summaries and Sales Mode Selection in Dual-Channel Supply Chains	Jiayao Zhao, Lin Wang, Zhiyong Li and Hongxin Liu (Communication Univ. of China)
	Conversational AI in High-Involvement Purchases: Chatbot-Assisted EV Buying and Channel Choice in Germany	Tianhao Li (Univ. of Göttingen)
	From Features to Semantics: Content Reconstruction with LLMs for Recommendation	Dahee Kim, Seonu Park (Kyung Hee Univ.), Qinglong Li (Hansung Univ.) and Sung-Byung Yang (Kyung Hee Univ.)
	Explaining E-Commerce Customer Engagement via Two-Part Modeling and LLM-Based Signal Analysis	Jiho Kim, Sangmin Park, Yooshin Kwon (Chung-Ang Univ.), Heongi Kim and Hyeonjae Cheon (GS Retail)

B 트랙

B1	**Smart Logistics & Supply Chain Resilience / Session Chair : Fen Qin (Taiyuan Univ. of Technology)**	
09:00 ~ 10:20	Sustainable Smart Logistics: A Systematic Review of Collaborative Mechanisms Enabled by Digital Technologies and the TOE Framework	Dongqing Huo (Anhui Univ. of Science and Technology), Jinxin Yang (Hefei Tech. College) and Chao Lu (Anhui Univ. of Science and Technology)
	How Does Information Technology Innovation Promote the Sustainable Development of Logistics Enterprises? – A Systematic Review Based on the TOE Framework	Hailin Cao (Anhui Univ. of Science and Technology), Jinxin Yang (Hefei Tech. College) and Chao Lu (Anhui Univ. of Science and Technology)
	AI Applications in Data Analysis and Automated Decision-Making in Logistics Smart Contracts	Thu Uyen Ha (Univ. of Hull)
	Mapping the Techno-Functional Architecture of Green Manufacturing: A Patent-based Analysis	Yu-Chao Cheng and Cheng-Hsien Hsieh (Singapore Univ. of Social Sciences)
	Blockchain-enabled Scope 3 Disclosure Strategies in Platform Supply Chains	Yuqian Wang (Renmin Univ. of China)
	From Engagement to Collapse: A Dynamic Change-Point Perspective on Gamified OMO Customer Experience	Su Jun Zhu and Jing Zhang (Kookmin Univ.)
	Digital Transformation, Resource Allocation and Agri-Supply Chain Resilience	Qianyi Wang and Ruyuan Huang (Shandong Technology and Business Univ.)
	Emotional Interaction and Conversion Value in Live Streaming E-commerce	Fen Qin (Taiyuan Univ. of Technology), Yang Li (Shandong Normal Univ.) and Miao Feng (Shandong Management Univ.)
B2	**Corporate Strategy & Competency / Session Chair : Jinxin Yang (Hefei Tech. College)**	
13:20 ~ 14:40	Resource Reconfiguration and Capability Regeneration: A Multilevel Mechanism Study on IT Innovation and Firm Performance	Chao Chu (Anhui Univ. of Science and Technology) and Jinxin Yang (Hefei Tech. College)
	Pragmatism over Greenery: Why Utilitarian Benefits Drive Smart Locker Adoption in Vietnam's Last-Mile Delivery	Minh-Trang Hoang, Chih-Hsiung Hu (National Formosa Univ.), Tien-Dat Nguyen (National Formosa Univ. & Van Hien Univ.) and Chih-Chin Liang (National Formosa Univ.)
	An Investigation into the Mechanisms of Enterprise Digital Transformation Through the Lens of New Institutional Theory	Yingzhao He and Chengchi Zhou (Northwest Normal Univ.)
	Interpretable Machine Learning for Predicting Employee Attrition in SMEs	Fotima Yusupova and So Ra Park (Chonnam National Univ.)
	Causal Inference of SEO Strategies: A Structural Equation Modeling Approach	Masaki Hayashi and Masamitsu Moriyama (Kindai Univ.)
	Digital Transformation and Frontline Employees' Adaptive Performance: A Moderated Mediation Model with Implications for AI in E-Commerce	Dandan Liu, Fengyi Wang, Kaiyuan Zhao and Xiongying Niu (Univ. of International Business and Economics)
	An Empirical Application of DIGO: An AI-Driven Distributed Commerce Platform with Organic Boosting	Jaehyung Son (Waplus)

B 트랙

B3	AI-Enabled E-Commerce & Consumer Decision / Session Chair : Jongwook Woo (California State Univ. Los Angeles)	
14:50 ~ 16:10	Nudging Sustainable Choices in Smart Commerce: AI Chatbots and the Personalization Paradox	Tien-Dat Nguyen, Tien-Dat Bui, Minh-Trang Hoang, and Chih-Chin Liang (National Formosa Univ.)
	Evaluating Secure Transaction Adoption and Third-Party Tracking Exposure in Modern E-Commerce Platforms	Waqar Javed (Atlantis Univ.)
	MindfulCart: AI Support for Cognitive Shopping	Nicholas Chung (California State Univ. Fullerton), Jennifer Jin (California State Univ. San Bernardino), Seunghyun Koo (Sungshin Women's Univ.) and Mira Kim (California State Univ. Fullerton)
	Scalable Sentiment Analysis for AI-Powered E-Commerce: Comparing ML and BERT-Based NLP Models on Large-Scale Amazon Book Reviews	Hsiu-Ping Lin California State Univ. Los Angeles), Sunnyoung Lee (Dongguk Univ.) and Jongwook Woo (California State Univ. Los Angeles)
	Multidimensional Rule-based Grocery Product Classification	Ruobing Liu and Hyunwoo Lim (Inha Univ.)
	Does Macroeconomic Uncertainty Drive Online Purchasing? Evidence from Daily Coupang Data	Jae Ik Ahn and Gyoo Gun Lim (Hanyang Univ.)
	Can Trust Signals Defend Against Negative Reviews? Evidence from the South Korean E-Commerce Market	Jae Ik Ahn, Seung Min Kim and Gyoo Gun Lim (Hanyang Univ.)
B4	**Smart Tourism & Immersive Experience / Session Chair : Soyeon Kwon (Dongguk Univ.)**	
16:20 ~ 17:40	Development of Hybrid Slot-Filling Travel Recommendation System: Focusing on the Runner-Pacer Dual Framework	Jae Ik Ahn, Su Kyeom Kim and Gyoo Gun Lim (Hanyang Univ.)
	Does Artificial Intelligence Enhance Tourism Economic Resilience? Evidence for Smart Tourism 2.0	Wanjun Gu and Chulmo Koo (Kyung Hee Univ.)
	Framing Scarcity: The Joint Effects of Availability Quantity and Presentation Format on Consumer Decisions	Danhye Kyoung and Chulmo Koo (Kyung Hee Univ.)
	Exploring the Actor-Network Formation of Edu-Metaverse: A Main Path Analysis	Van-Tan Nguyen, Chen-Hao Huang and Ching-Ying Huang (National Cheng Kung Univ.)
	Generation Z Generative AI Travel Decision-Making Value Co-Creation	Ganghao Zhang, Xinyuan Xu, Mengqiang Ma, Yuxin Li, Jingjing Liu, Fuying Xu, Tongzheng Wang, Safa Ait Larbi (Shandong Technology and Business Univ.) and Heesup Han (Sejong Univ.)
	Extending Perceived Risk Framework in Live Streaming E-Commerce: An Information Ecology Perspective	Zitong Li, Zhaohong Liang and Xiaosong Wu (Beijing Normal Univ.)

C 트랙

C1	**Responsible AI & Digital Governance / Session Chair : Wenting Liu (Singapore Univ. of Social Sciences)**	
09:00 ~ 10:20	Exploring Ethical Challenges of Artificial Intelligence in Commercial Contexts	Ling-Pu Huang and Chen-Hao Huang (National Cheng Kung Univ.)
	How AI Transparency Affects AI Appreciation: The Mediating Roles of Trusting Beliefs	Hua Liu (Anhui Univ.), Xiaofan Tang (Anhui Medical Univ.) and Jinmei Yin (Nanjing Univ. of Aeronautics and Astronautics)
	AI-Driven Intelligent Policy Monitoring Method: Measuring the Intensity and Synergy of China's Data Element Market	Erqiao Lin, Xubu Ma and Chunxiu Qin (Xidian Univ.)
	Perception of AI Bureaucracy on a Value-Rational and Public Communication Proposals	Jinsoon Song (Donga Univ.)
	Information Sensitivity in the AI-Driven Digital Society: Constructs and Measurement Instrument	Xi Chen, Hongying Du (KMMU), Jian Mou (Pusan National Univ.), and Pengxin Zheng (Yunnan Univ.)
	When Citizens Become Consumers of AI: Algorithm Aversion, Appreciation and Compliance Behavior in AI-Driven Public Health Warning Services	Hsiao-Ting Tseng (National Central Univ.)
	An Exploratory Framework for Flagging Greenwashing: Combining NLP-Based Patent Analysis with Financial-Patent Efficiency Metrics	Wen Yan Toh, Wenting Liu (Singapore Univ. of Social Sciences), and Nat Liang (Independent Researcher)
C2	**Digital Media, Advertising & Content / Session Chair : Jee-Hae Lim (Univ. of Hawaii at Mānoa)**	
	Who Tells It Better? Aligning Creation Modalities with Types of PSAs to Enhance Public Attitudes via Systematic Thinking	Xiaoxiao Gong and Danfeng Yuan (GuiZhou Univ.)
	Can Machines Read Your Face? A Video Analytics Framework for Measuring Empathy via Emotional Mimicry in Video Recordings	Li Cui (The Hong Kong Polytechnic Univ.), Jee-Hae Lim (Univ. of Hawaii at Manoa), Zhoudao Lu (Univ. of Oxford), Ka Chung Ng, and Jingran Zhao (The Hong Kong Polytechnic Univ.)
	A Cost-Sensitive Information Grading Model Considering Asymmetric Misclassification Risks	Seojin Lee, Yeji Choi, Wonyoung Kim, Hyeseong Lee and Hangbae Chang (Chung-Ang Univ.)
	How AIGC Disclosure Labels Drive Consumers' UGC Compensatory Search Behavior	Zhenhan Qin, Xusen Cheng, Xinhe Qu, Ang Zeng and Yijun Yan (Renmin Univ. of China)
	Analysis of the Influences of AI on the Media Industry with Story Telling	Hui Liu and Jinhwa Kim (Sogang Univ.)
	Intrusive or Effective? The Trade-Off between Brand Recall and Viewer Satisfaction of In-Game Advertising in eSports	Sze Ming Chow, Yan Sun and Sung-Byung Yang (Kyung Hee Univ.)

C 트랙

C3	**Digital Finance & Platform Economy / Session Chair : Kyung Young Lee (Dalhousie Univ.)**	
14:50 ~ 16:10	Understanding the Endorsement Effect of Digital Virtual Human on Investment Intention Towards Crowdfunding-Based on Elaboration Likelihood Model	Jiajie Chu (Southwestern Univ. of Finance and Economics)
	Probabilistic Earnings Forecasting Using Large Language Models in Digital Platforms	Xiaolin Gan, Xincan Qiu and Chenglong Zhang (Kookmin Univ.)
	Analysis of Solana Validator Staking Amount: A Machine Learning Approach	Ke Li and Chenglong Zhang (Kookmin Univ.)
	Service Fee in C2C Secondhand Markets: Optimal Fee Design for Platform Governance Value	Guojun Ji, Chang Liu(Xiamen Univ.), and Kim Hua Tan (Nottingham Univ)
	How AI Credit Scoring Influences Consumer Satisfaction and Purchase Intention in AI-Driven E-Commerce: Evidence from Bibli PayLater	Sheila Silvia Permatasari and Syuan Yi Jiang (National Cheng Kung Univ.)
	Does the Adoption of Digital Wealth Management Platforms (DWPs) Result in Positive Net Benefits? Retail Investors' Perspective on DWP Success	Yong Yin, Kyung Young Lee, Bo Yu, Colin Conrad and Thomas Kb Koo (Dalhousie Univ.)
	Generative AI-Enabled Value Creation in Robo-Advisory: A Grounded Theory Multiple-Case Study	Qin Gao, Weihui Li and Yongqing Yang (Shandong Technology and Business Univ.)
C4	**Human Interaction & Future Intelligent Environments / Session Chair : Wenlong Zhu (Qingdao Univ. of Technology)**	
16:20 ~ 17:40	How Does Human-AI Fit Shape Consumers' Passive Decision-Making?	Wenlong Zhu and Lei Liu (Qingdao Univ. of Technology)
	Building Trust and Ties: Why Elderly Patients Choose Online Health Communities	Jinxin Yang (Hefei Tech. College) and Xinyi Xie (King's College of London)
	BhashaBridge: An Emotion-Aware Real-Time Indian Sign Language Translator with Offline Doctor Mode Using MediaPipe and LSTM	Er.Gautam Kumar, Surbhi Kumari, Roshni Soni, Isha Rani and Manavjeet Singh (Chandigarh Univ.)
	A Novel Lightweight Medical Image Encryption Scheme Using Convolutional Autoencoders and Chaotic Systems	Shaik. Kashif Hussain, Pedaballe Lekha Reddy, Kuruba Kullai Swamy and V. Kowshik Chowdary (Rajeev Gandhi College)
	Experiencing Joint Agency and Health Self-Efficacy in AI Healthcare: How Perceived Agency Influences Continuance Intention Under Two Dimensions of Anthropomorphism	Xiaoan Li and Jing Zhang (Kookmin Univ.)
	A Story Mining Approach to Solve Problems: A Case on Bridging Metaverse and Elderly	Hyunsung Jung, Soyeon Lee and Jinhwa Kim (Sogang Univ.)
	Research on the Impact of Human-AI Interaction Types on Individual Innovation Performance: A Dual-Pathway Model of Responsibility Attribution and Identification	Yujuan Zheng, Shuangwu Dong (XIdian Univ.) and Hanzhuo Ma (Deakin Univ.)
	Permission Structures, Inferential Processing and Perceived Power Asymmetry in Humanoid Home Robot Contexts	Jinjing Li, Jingshi Jin (Kyung Hee Univ.), Kyung Young Lee (Dalhousie Univ.) and Sung-Byung Yang (Kyung Hee Univ.)
	Embodied AI and Environmeoid Robots Reshape Privacy Calculus in Home Environments	Jingshi Jin, Jinjing Li and Sung-Byung Yang (Kyung Hee Univ.)

KITS·ICEC·KIISS
2026 통합 국제학술대회

W1 Live Streaming & Social Media Engagement

Session Chair : Dongyeon Kim (Hanyang Univ.)



W1

To narrate or to list? The interplay of avatar type, presentation style, and time-of-day effect in virtual livestreaming interaction.

Xiaoxiao Gong (Guizhou University), Jinhong Wang (Guizhou University)

*교신저자

This study investigates how aligning AI-driven virtual streamers' avatar types with presentation styles impacts consumer outcomes in e-commerce. Drawing on processing fluency theory and employing a mixed-method design, including text mining of 17,376 comments and four experiments (N=852). We examine the interaction between avatar type (Human Avatar Virtual Streamers [HAVSs] vs. Non-human Avatar Virtual Streamers [NAVSs]) and presentation style (Narrative vs. List). Results reveal consumer responses are maximized when HAVSs employ a narrative style, which enhances affective fluency, and when NAVSs utilize a list style, which enhances cognitive fluency. Furthermore, we identify time-of-day as a psychophysiological boundary condition regulating this mechanism.The HAVS-Narrative effect is amplified at night when users favor affective fluency, while the NAVS-List match excels during the day when cognitive resources support cognitive fluency. Ultimately, synchronizing avatar types and presentation styles with users' time-dependent cognitive states optimizes AI-mediated interactions and drives favorable consumer evaluations.

KEY WORDS

Avatar type, Presentation style, Affective fluency, Cognitive fluency, Time-of-day effect

W1

Understanding Paid Subscription Behavior in YouTube Memberships: The Role of Emotional Attachment.

Yen-Hao Hsieh, Ru-Yi Lai, Chih-Wei Hsu, Pei-Syuan Jiang, Pei-Qi Wu, Yi-Ching Hsieh, Zheng-Zhe Chen
(National Taichung University of Science and Technology)

This study investigates paid subscription behavior in YouTube memberships by examining the roles of parasocial interaction, perceived value, substitutability, and emotional attachment. Using survey data collected from YouTube membership subscribers and analyzing the data with PLS–SEM, the results show that parasocial interaction significantly enhances emotional attachment, which in turn strongly predicts paid subscription intention. In contrast, perceived value and content substitutability do not have significant effects on emotional attachment. These findings suggest that paid subscription behavior in digital video platforms is primarily driven by relational and emotional mechanisms rather than rational evaluations of value or alternative content options. By clarifying the psychological pathway from parasocial interaction to financial commitment, this study contributes to the understanding of subscription behavior in online content platforms and provides practical insights for creators seeking to strengthen audience engagement and membership sustainability.

KEY WORDS

Paid subscription, Emotional attachment, Parasocial interaction, Digital platforms, YouTube memberships

Understanding Paid Subscription Behavior in YouTube Memberships: The Role of Emotional Attachment

Hsieh, Y.-H.[a] , Lai, R.-Y.[b], Hsu, C.-W.[c], Jiang, P.-S.[d], Wu, P.-Q.[e], Hsieh, Y.-C.[f], and Chen, Z.-Z.[g]

[a] Department of Business Administration, National Taichung University of Science and Technology
E-mail: yhhsiehs@nutc.edu.tw
[b] Department of Business Administration, National Taichung University of Science and Technology
E-mail: zoya930201@gmail.com
[c] Department of Business Administration, National Taichung University of Science and Technology
E-mail: hsu921023@gmail.com
[d] Department of Business Administration, National Taichung University of Science and Technology
E-mail: patty93523@gmail.com
[e] Department of Business Administration, National Taichung University of Science and Technology
E-mail: Peggy9304271230@gmail.com
[f] Department of Business Administration, National Taichung University of Science and Technology
E-mail: a0965652777@gmail.com
[g] Department of Business Administration, National Taichung University of Science and Technology
E-mail: as0983924583@gmail.com

Abstract

This study investigates paid subscription behavior in YouTube memberships by examining the roles of parasocial interaction, perceived value, substitutability, and emotional attachment. Using survey data collected from YouTube membership subscribers and analyzing the data with PLS-SEM, the results show that parasocial interaction significantly enhances emotional attachment, which in turn strongly predicts paid subscription intention. In contrast, perceived value and content substitutability do not have significant effects on emotional attachment. These findings suggest that paid subscription behavior in digital video platforms is primarily driven by relational and emotional mechanisms rather than rational evaluations of value or alternative content options. By clarifying the psychological pathway from parasocial interaction to financial commitment, this study contributes to the understanding of subscription behavior in online content platforms and provides practical insights for creators seeking to strengthen audience engagement and membership sustainability.

Keywords: Paid subscription; Emotional attachment; Parasocial interaction; Digital platforms; YouTube memberships

Introduction

In recent years, YouTube has become one of the world's leading digital video platforms, reporting approximately 2.7 billion active users globally and over one billion hours of video viewed daily. In Taiwan, the platform reaches more than 80% of the population, underscoring its dominant role in digital media consumption. Alongside its rapid growth, the creator economy has evolved from advertising-based revenue toward diversified subscription-oriented mechanisms, including channel memberships and interactive payment features (Li et al., 2022). Despite these monetization tools, motivating viewers to cross the payment threshold remains a persistent challenge in an environment characterized by abundant free content and low switching costs.

Prior research suggests that subscription decisions are shaped not only by content quality but also by viewers' psychological experiences (Chen, 2023). High quality experience can enhance perceived value, while the intangible and irreversible nature of digital purchases heightens perceived risk. Moreover, when alternative content is readily available, audiences may hesitate to commit financially. In such contexts, emotional engagement becomes critical. Through parasocial interaction, creators may cultivate emotional attachment that transforms subscription from a transactional act into a relational commitment (Eom et al., 2025).

Although existing studies have examined psychological or contextual factors independently, limited research has integrated these mechanisms into a unified framework explaining paid subscription behavior (Sokolova and Kefi, 2020; Zhang and Zhang, 2024). Accordingly, this study investigates how parasocial interaction, perceived value, and emotional attachment jointly influence YouTube viewers' paid subscription behavior under conditions of content substitutability. By developing and empirically testing an integrative decision-making model, this research aims to clarify the psychological processes underlying subscription behavior and provide managerial insights for optimizing membership strategies in digital video platforms.

Hypotheses Development

Perceived Value and Emotional Attachment

In digital content consumption contexts, viewers' perceived value of content, encompassing both functional benefits and emotional gratification, may influence their overall evaluation of the content source (Oyedele and Simpson, 2018). When individuals perceive that the content they engage with provides tangible benefits and positive experiences, they are more likely to develop favorable attitudes toward the creator (Sweeney and Soutar, 2001). Through repeated exposure and ongoing interaction, such positive evaluations may gradually evolve into more stable psychological connections.

From a relational development perspective, perceived value can be considered a foundational condition for relationship formation (Wang et al., 2013). When viewers recognize the overall value delivered by a creator, they are more likely to develop identification and, over time, form emotional attachment (Chen and Chen, 2010). Accordingly, in the context of online video consumption, higher perceived value may facilitate the development of emotional attachment toward the content creator. Based on the above reasoning, this study proposes the following hypothesis:

H1: Perceived value has a significant positive effect on emotional attachment.

Substitutability and Emotional Attachment

In digital content markets, viewers are exposed to a wide range of alternative options. Substitutability reflects the extent to which individuals perceive the availability of comparable alternatives. When multiple functionally similar and easily accessible content sources are available, viewers' motivation to develop long-term emotional bonds with a single creator may be affected (Jones et al., 2000).

From a relational investment perspective, the formation of emotional attachment typically involves the allocation of time and psychological resources (Burnham et al., 2003). If individuals perceive that alternative options are abundant and switching costs are low, they may be less inclined to invest deeply

in a particular relational tie (Ping, 1993). In highly substitutable environments, viewers may prefer to maintain flexible consumption patterns rather than develop stable and exclusive emotional connections with a specific creator.

Accordingly, in the context of online video platforms, higher perceived content substitutability may be associated with lower levels of emotional attachment toward the creator. Based on the above reasoning, this study proposes the following hypothesis:

H2: Substitutability has a significant negative effect on emotional attachment.

Parasocial Interaction and Emotional Attachment

In digital media contexts, parasocial interaction refers to viewers' subjective perception of experiencing a face-to-face–like interaction with media personalities. When viewers perceive that a creator responds to them, understands them, or conveys emotions during the viewing process, such experiences may generate a sense of relational connection, even though the interaction is inherently one-sided.

From a relational development perspective, repeated interaction experiences can strengthen feelings of trust and identification, thereby encouraging emotional investment (Labrecque, 2014). When viewers consistently perceive an ongoing and meaningful interaction with a creator, they are more likely to internalize this perceived connection and develop emotional attachment (Tukachinsky, 2011). In this sense, parasocial interaction functions as an important relational mechanism through which repeated exposure and perceived psychological closeness foster the formation of emotional bonds (Sun and Li, 2026).

Accordingly, in the context of online video platforms, stronger perceived parasocial interaction is expected to increase viewers' emotional attachment toward the creator. Based on the above reasoning, this study proposes the following hypothesis:

H3: Parasocial interaction has a significant positive effect on emotional attachment.

Emotional Attachment and Paid Subscription Behavior

Emotional attachment is commonly regarded as a fundamental psychological basis for relationship maintenance and supportive behaviors (Koskie and Locander, 2023). When individuals develop a stable and enduring emotional bond with a particular target, they are more likely to feel motivated to sustain that relationship and to invest time, attention, and resources to maintain it (Pang and Zhang, 2024). Emotional attachment reflects not only affective closeness but also a subjective sense of relational value (Novak et al., 2000).

In the context of online video platforms, viewers' emotional attachment to a creator may translate into concrete supportive behaviors. Compared with passive viewing, paid subscription represents a form of tangible commitment, as viewers are willing to incur financial costs to express identification with and support for the creator. Therefore, when viewers experience stronger emotional attachment toward a creator, they are more likely to maintain the relationship through paid subscription (Kim et al., 2013).

In this sense, emotional attachment can be viewed as a key mechanism that links psychological identification with actual behavioral outcomes, transforming affective bonds into concrete financial support. Based on the above reasoning, this study proposes the following hypothesis:

H4: Emotional attachment has a significant positive effect on paid subscription behavior.

Research Model

This study proposes an integrative model linking antecedent factors, emotional attachment, and paid subscription behavior (as shown in Figure 1).

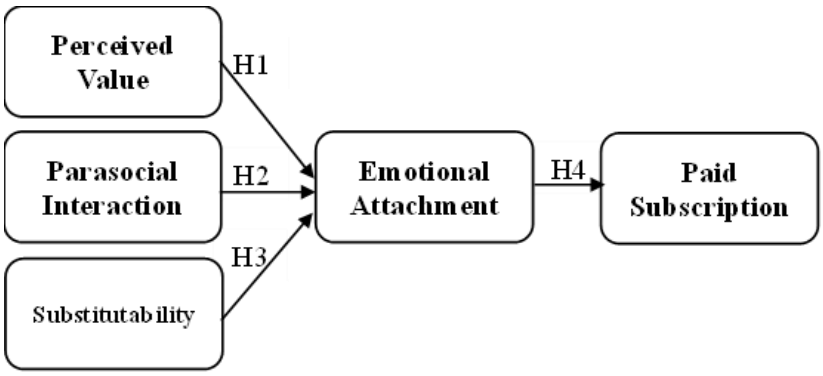


Figure 1. Research Model

Methodology

Questionnaire Design

The measurement items used in this study were adapted from previously validated scales reported in the relevant literature, all of which demonstrated satisfactory reliability and validity. To ensure contextual appropriateness, minor modifications were made to align the wording of the items with the research setting. All constructs were measured using a seven-point Likert scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"), allowing respondents to indicate their perceptions and evaluations of each construct.

Table 1. Demographic Description

Variable	**Item**	**Percentage (%)**
Gender	Male	33.7
	Female	64.4
	Decline to answer	2
Age	18-20	12.9
	21-30	70.3
	31-40	5.9
	> 40	10.9
Education	High school	8.9
	University	73.3
	Postgraduate	17.8
Usage frequency (per week)	Less than 3 times	59.4
	4-10 times	20.8
	More than 10 times	19.8
Average daily viewing time	< 1 hour	59.4
	1 – 3 hours	31.7
	> 3 hours	8.9

Data Collection

This study employed a survey method to collect empirical data. The target respondents were individuals who had previously paid for YouTuber channel memberships, ensuring that participants possessed actual subscription experience. Data were collected through an online questionnaire administered via Google Forms. The survey link was distributed across various social networking platforms to increase accessibility and broaden the sample pool.

Data Analysis

SmartPLS 4 was employed to test the proposed hypotheses using structural equation modeling (SEM). This analytical approach enables the simultaneous examination of relationships among multiple latent and observed variables and is particularly suitable for evaluating complex causal models. The analysis proceeded in several stages. First, descriptive statistical analyses were conducted to summarize the demographic characteristics of the sample. Second, reliability and validity assessments were performed to evaluate the measurement quality of the constructs. Finally, structural model analysis was conducted to examine the hypothesized relationships and assess the overall model fit and path relationships among constructs.

Table 2. Reliability and Validity

Construct	Item	Factor Loadings	Cronbach's Alpha	CR	AVE
PV	PV1	0.800	0.881	0.905	0.577
	PV2	0.733			
	PV3	0.782			
	PV4	0.815			
	PV5	0.724			
PSI	PSI1	0.776	0.899	0.920	0.624
	PSI2	0.820			
	PSI3	0.801			
	PSI4	0.822			
	PSI5	0.763			
	PSI6	0.800			
SUB	SUB1	0.864	0.938	0.948	0.722
	SUB2	0.762			
	SUB3	0.912			
	SUB4	0.904			
	SUB5	0.878			
	SUB6	0.834			
EA	EA1	0.863	0.943	0.954	0.747
	EA2	0.869			
	EA3	0.875			
	EA4	0.838			
	EA5	0.863			
	EA6	0.847			
PSUB	PSUB1	0.791	0.932	0.946	0.715
	PSUB2	0.849			
	PSUB3	0.907			
	PSUB4	0.878			
	PSUB5	0.902			

Statistical Analysis and Results

Sample Characteristics

A total of 131 questionnaires were collected. After excluding respondents who had not subscribed to any YouTuber channel membership, 30 invalid responses were removed, resulting in 101 valid questionnaires for subsequent analysis. As shown in Table 1, female respondents constituted the majority of the sample (64.4%). The age distribution was primarily concentrated in the 21–30 age group. Most respondents held a bachelor's degree (73.3%). Regarding monthly income, the largest proportion of participants reported earning below NT$20,000 (50.5%). With respect to usage behavior, 20.8% of

respondents reported watching YouTube four to ten times per week. Additionally, 41.6% indicated that they had paid to watch YouTuber membership-exclusive videos within the past month.

Measurement Model

Prior to testing the structural relationships, the reliability and validity of the measurement model were examined. All constructs demonstrated satisfactory internal consistency, with Cronbach's alpha values ranging from 0.881 to 0.952 and composite reliability (CR) values exceeding the recommended threshold of 0.70 (as shown in Table 2).

Convergent validity was supported, as the average variance extracted (AVE) values ranged from 0.577 to 0.748, exceeding the minimum criterion of 0.50. Discriminant validity was confirmed using the Fornell–Larcker criterion and the HTMT ratio, indicating that the constructs were empirically distinct. Overall, the measurement model met the established standards for reliability and validity (as shown in Table 3 and Table 4).

Table 3. Fornell–Larcker

Construct	EA	PSI	PSUB	PV	SUB
EA	0.864				
PSI	0.747	0.790			
PSUB	0.704	0.631	0.845		
PV	0.650	0.771	0.594	0.760	
SUB	0.266	0.383	0.191	0.211	0.850

Structural Model

The structural model was examined using PLS-SEM with bootstrapping (5,000 resamples) to evaluate the significance of the hypothesized relationships. The standardized root mean square residual (SRMR) was 0.089, indicating an acceptable level of model fit based on commonly used guidelines. The results show that parasocial interaction had a significant positive effect on emotional attachment ($\beta = 0.351$, $t = 2.591$, $p = 0.010$). In turn, emotional attachment had a significant positive effect on paid subscription intention ($\beta = 0.704$, $t = 10.439$, $p < 0.001$). In contrast, perceived value ($\beta = 0.054$, $t = 0.383$, $p = 0.702$) and substitutability ($\beta = 0.003$, $t = 0.035$, $p = 0.972$) did not have significant effects on emotional attachment. Overall, the structural results indicate that the formation of paid subscription intention is primarily associated with the relational pathway in the proposed model, with parasocial interaction fostering emotional attachment, which subsequently increases viewers' paid subscription intention.

Table 4. HTMT

Construct	EA	PSI	PSUB	PV
PSI	0.801			
PSUB	0.744	0.629		
PV	0.667	0.187	0.224	
SUB	0.244	0.063	0.158	0.155

Mediation Analysis

To further examine how antecedent factors translate into paid subscription intention, this study tested the mediating role of emotional attachment between the antecedent variables and paid subscription intention. A bootstrapping procedure with 5,000 resamples was used to estimate specific indirect effects. Mediation significance was assessed based on bias-corrected confidence intervals and associated p-values.

As shown in Table 6, parasocial interaction exerted a significant indirect effect on paid subscription intention through emotional attachment ($\beta = 0.247$, $p = 0.014$). This result indicates that viewers' perceived interaction with creators enhances emotional attachment, which subsequently contributes to higher subscription intention.

In contrast, the indirect effects of perceived value and substitutability on paid subscription intention via emotional attachment were not statistically significant (PV: $\beta = 0.038$, $p = 0.604$; SUB: $\beta = 0.002$, $p = 0.975$). These findings suggest that, within this context, cognitive evaluations and perceptions of alternative availability did not translate into subscription intention through the attachment mechanism.

Table 5. Structural Results

Hypothesis	β	t-Value	p-Value	Support
H1	0.054	0.383	0.702	No
H2	0.003	0.035	0.972	No
H3	0.351	2.591	0.010	Yes
H4	0.704	10.439	<0.001	Yes

Conclusion and Implications

Conclusion

This study examined the factors influencing paid subscription behavior in YouTube membership programs by integrating parasocial interaction, perceived value, substitutability, and emotional attachment into a unified analytical framework. The empirical results reveal that parasocial interaction significantly enhances viewers' emotional attachment to content creators, which in turn strongly predicts paid subscription intention. In contrast, perceived value and content substitutability do not significantly influence emotional attachment. These findings indicate that subscription behavior in digital content platforms is shaped primarily by relational and emotional mechanisms rather than by purely rational evaluations of content value or comparisons with alternative options.

The results contribute to the literature on digital platform monetization by highlighting the critical role of emotional attachment in transforming audience engagement into financial commitment. Specifically, the study demonstrates that parasocial interaction functions as a key relational pathway through which viewers develop emotional bonds with creators, ultimately encouraging paid support. From a managerial perspective, creators should prioritize strategies that strengthen audience interaction and relational engagement rather than focusing solely on increasing informational or functional value in their content offerings.

Discussion

First, the findings indicate that perceived value did not exert a statistically significant effect on emotional attachment. While prior research has emphasized the importance of perceived value in shaping behavioral intentions, the present results suggest that perceived value alone may not be sufficient to foster emotional attachment in the YouTube membership context. Although viewers may

recognize the functional or informational benefits of content, such evaluations do not automatically translate into deeper emotional bonds. Emotional attachment may require sustained interaction and relational engagement beyond cognitive value assessment.

Second, the results show that content substitutability did not significantly influence emotional attachment. This finding suggests that, within the current sample and context, viewers' emotional bonds with creators were not directly affected by perceptions of alternative content availability. Even in environments characterized by abundant free alternatives, emotional attachment may not be primarily determined by comparative evaluation processes. Instead, relational and experiential factors appear to play a more direct role in shaping attachment formation.

Managerial Implications

The results indicate that, compared to rational value evaluations, parasocial interaction and emotional attachment play more central roles in shaping paid subscription intention. This finding suggests that creators should not position themselves merely as content providers but should actively cultivate relational engagement with their audiences. Real-time interaction, members-only communication channels, and the sharing of personal or behind-the-scenes content may help strengthen viewers' emotional bonds and, consequently, enhance subscription intention.

Although perceived value did not directly influence emotional attachment, it may still function as a foundational condition for sustaining long-term membership retention. Exclusive content in membership areas should avoid substantial overlap with publicly available materials. In addition, symbolic benefits, such as member badges or identity markers, may enhance the perceived meaning of subscription behavior. By combining emotional connection with appropriate value reinforcement, creators may foster more stable and enduring subscription relationships within the YouTube membership ecosystem.

Limitations and Future Research

This study has several limitations. First, the sample was limited to YouTube membership users, which may restrict the generalizability of the findings to other digital platforms. Second, the cross-sectional design limits the ability to infer causal relationships. Future research may employ longitudinal or experimental approaches to further examine the dynamic relationship between parasocial interaction, emotional attachment, and subscription behavior across different digital platform contexts.

References

Burnham, T. A., Frels, J. K., and Mahajan, V. 2003. "Consumer Switching Costs: A Typology, Antecedents, and Consequences," *Journal of the Academy of Marketing Science* (31:2), pp. 109-126.

Chen, L. 2023. "Analysis of Online Platforms' Free Trial Strategies for Digital Content Subscription," *Journal of Theoretical and Applied Electronic Commerce Research* (18:4), pp. 2107-2124.

Chen, C. F., and Chen, F. S. 2010. "Experience Quality, Perceived Value, Satisfaction and Behavioral Intentions for Heritage Tourists," *Tourism Management* (31:1), pp. 29-35.

Eom, S., Park, J., Choi, E., Park, J., and Kim, S. 2025. "How Do Users Contribute to YouTube Channels' Revenue? An Empirical Analysis of Korean Beauty Channels," *Computers in Human Behavior* (172), Article 108741.

Hoffman, D. L., and Novak, T. P. 1996. "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing* (60:3), pp. 50-68.

Jones, M. A., Mothersbaugh, D. L., and Beatty, S. E. 2000. "Switching Barriers and Repurchase Intentions in Services," *Journal of Retailing* (76:2), pp. 259-274.

Kim, J., Kim, M., and Wachter, K. 2013. "A Study of Mobile User Engagement and Behavioral Intentions," *Journal of Computer Information Systems* (53:3), pp. 82-90.

Koskie, M. M., and Locander, W. B. 2023. "Cool Brands and Hot Attachments: Their Effect on Consumers' Willingness to Pay More," *European Journal of Marketing* (57:4), pp. 905-929.

Labrecque, L. I. 2014. "Fostering Consumer–Brand Relationships in Social Media Environments," *Journal of Interactive Marketing* (28:2), pp. 134-148.

Li, Z., Wang, D., Nan, G., and Li, M. 2022. "Optimal Revenue Model of a Social Networking Service: Ad-Sponsored, Subscription-Based, or Hybrid?," *IEEE Transactions on Engineering Management* (71), pp. 1-13.

Novak, T. P., Hoffman, D. L., and Yung, Y. F. 2000. "Measuring the Customer Experience in Online Environments: A Structural Modeling Approach," *Marketing Science* (19:1), pp. 22-42.

Oyedele, A., and Simpson, P. M. 2018. "Streaming Apps: What Consumers Value," *Journal of Retailing and Consumer Services* (41), pp. 296-304.

Pang, H., and Zhang, K. 2024. "Determining Influence of Service Quality on User Identification, Belongingness, and Satisfaction on Mobile Social Media: Insight from Emotional Attachment Perspective," *Journal of Retailing and Consumer Services* (77), Article 103688.

Ping, R. A. 1993. "The Effects of Satisfaction and Structural Constraints on Retailer Exiting, Voice, Loyalty, Opportunism, and Neglect," *Journal of Retailing* (69:3), pp. 320-352.

Sokolova, K., and Kefi, H. 2020. "Instagram and YouTube Bloggers Promote It, Why Should I Buy? How Credibility and Parasocial Interaction Influence Purchase Intentions," *Journal of Retailing and Consumer Services* (53), Article 101742.

Sun, X., and Li, J. 2026. "Parasocial Bonds Drive Brand Connection: How Listener–Host Parasocial Relationships Enhance Brand Advocacy through Emotional Investment," *Journal of Retailing and Consumer Services* (90), Article 104712.

Sweeney, J. C., and Soutar, G. N. 2001. "Consumer Perceived Value: The Development of a Multiple Item Scale," *Journal of Retailing* (77:2), pp. 203-220.

Tukachinsky, R. 2011. "Parasocial Relationships: Development of a Multidimensional Measure," *Communication Research Reports* (28:2), pp. 113-123.

Wang, Y.-S., Yeh, C.-H., and Liao, Y.-W. 2013. "What Drives Purchase Intention in the Context of Online Content Services? The Moderating Role of Ethical Self-Efficacy for Online Piracy," *International Journal of Information Management* (33:1), pp. 199-208.

Zhang, L., and Zhang, Y. 2024. "Whether to Add a Digital Product into Subscription Service?," *Journal of Theoretical and Applied Electronic Commerce Research* (19), pp. 45-58.

W1

Exploring the Thematic Structure of Live Streaming Commerce: A Bibliometric Analysis using LDA.

Yu-Shan Huang, Chen-Hao Huang (National Cheng Kung University)

Live streaming commerce (LSC) has rapidly emerged as a transformative force in the online market by facilitating real-time interactions between consumers and streamers. This phenomenon has attracted scholarly attention across various disciplines. However, despite the growing body of research, literature remains fragmented across different themes, making it difficult to develop a coherent understanding of the field and providing limited guidance for businesses in establishing effective differentiation strategies. To address this gap, this study conducted a systematic review that combined bibliometric analysis and topic modeling to examine the thematic structure of LSC. The preliminary findings identified three major thematic clusters: commercial-oriented research, experience-oriented research, and technology-oriented research. By exploring these themes, this study enhances the understanding of LSC and its implications for the future.

KEY WORDS

Live Streaming Commerce, Main Path Analysis, Latent Dirichlet Allocation

W1

AI-Inferred Personality Traits and Humanlikeness in Software Reviews: Evidence from Collaboration Tool Platforms.

Sunnyoung Lee (Dongguk University),
Hsiu-Ping Lin, Jongwook Woo (California State University Los Angeles)

This study examines how personality traits inferred through artificial intelligence based text analysis influence user evaluations of collaborative software. Using approximately 18,000 user reviews from G2.com for Notion and ClickUp, the study estimates Big Five personality traits from review texts and analyzes their relationship with user ratings. Perceived humanlikeness of the software is incorporated as a moderating factor, while sentiment score is included as a control variable. The results indicate that review sentiment is the strongest predictor of ratings across both platforms. The direct effects of personality traits are generally limited, but several interaction patterns suggest that perceived humanlikeness may shape how personality traits relate to user evaluations. These findings provide insight into the role of psychological characteristics and AI related perceptions in the evaluation of digital collaboration tools.

KEY WORDS

Artificial intelligence, Big Five personality, Humanlikeness

W1

When Interaction Matters: Learning Effect of Knowledge-Sharing Live Streaming.

SHanqiong Wu, Zhengzhi Guan (Beijing Normal - Hong Kong Baptist University),
Bosen Lin (University of Nottingham Ningbo China).

This study examines live streaming as an emerging trend in knowledge-sharing contexts, where interaction is recognized as a distinctive feature but lacks systematic frameworks for effective learning implementation. While prior research has largely focused on viewer behaviour in entertainment and e-commerce live streaming, limited attention has been paid to how interaction influences knowledge learning effects in such settings. Using a two-phase mixed-method approach, this study combines survey data analyzed via PLS-SEM with supplementary focus group insights. The results show that social presence, synchronicity, and active control significantly enhance knowledge-sharing effects, with social presence exerting the strongest influence, while two-way communication is not significant overall. Qualitative findings further provide plausible explanation for quantitative findings. By focusing on knowledge-sharing effects in interactive live streaming environments, this study develops and validates key factors that explain how interaction shapes effective knowledge learning.

KEY WORDS

Knowledge-Sharing Live Streaming, Learning Effects, Interactivity, Social Presence

When Interaction Matters: Learning Effect of Knowledge-Sharing Live Streaming

Hanqiong Wu [a], Zhengzhi Guan [b] and Bosen Lin [c]
[a] Faculty of Business and Management, Beijing Normal-Hong Kong Baptist University
E-mail: hanzel_w@163.com
[b] Faculty of Business and Management, Beijing Normal-Hong Kong Baptist University
E-mail: zhengzhiguan@bnbu.edu.cn
[c] Nottingham University Business School China
E-mail: bosen.lin@nottingham.edu.cn

Abstract

This study examines live streaming as an emerging trend in knowledge-sharing contexts, where interaction is recognized as a distinctive feature but lacks systematic frameworks for effective learning implementation. While prior research has largely focused on viewer behaviour in entertainment and e-commerce live streaming, limited attention has been paid to how interaction influences knowledge learning effects in such settings. Using a two-phase mixed-method approach, this study combines survey data analyzed via PLS-SEM with supplementary focus group insights. The results show that social presence, synchronicity, and active control significantly enhance knowledge-sharing effects, with social presence exerting the strongest influence, while two-way communication is not significant overall. Qualitative findings further provide plausible explanation for quantitative findings. By focusing on knowledge-sharing effects in interactive live streaming environments, this study develops and validates key factors that explain how interaction shapes effective knowledge learning.

Keywords: Knowledge-Sharing Live Streaming; Learning Effects; Interactivity; Social Presence

Introduction

Live streaming has become a major form of real-time digital media, with its market size expected to grow rapidly in the coming years (Mordor Intelligence 2026). Beyond entertainment and e-commerce, it is increasingly used for communication and knowledge sharing (360iResearch 2026). This reflects a broader shift in how people learn and share knowledge in digital environments.

With the rise of social media, knowledge learning has become more decentralized and interactive, supported by platforms such as blogs, wikis, and online communities. Recently, knowledge-sharing live streaming has emerged as a new form of knowledge exchange, characterized by real-time interaction (Yang et al. 2022). Since around 2018, this format has gained increasing attention as a distinct type of live streaming (Lu et al. 2018). It allows creators to share knowledge while interacting with audiences in real time, offering a more engaging experience than traditional online learning and also helps make knowledge more accessible to a wider audience (Fonseca and Barbosa 2021).

Prior studies show that interactivity can increase engagement, trust, and participation in live streaming (Hou et al. 2024). Social presence, additionally, enhances users' sense of connection in online environments (Hou et al. 2025). This indicate that the social interaction, characterized by interactivity and social presence, plays a critical role in live streaming. However, most existing studies focus on commercial outcomes across various live streaming contexts. We still know little about whether live streaming interaction improves knowledge learning effects, especially given its unique interactive

features. This gap is important because learning practices differ from commercial and entertainment activities. Therefore, this study examines how core interaction affordances of live streaming—two-way communication, active control, synchronicity, and social presence—affect knowledge learning effects in knowledge-sharing live streaming. By shifting the focus from commercial outcomes to learning outcomes, this study extends prior research and provides new insights into knowledge sharing, aiming to answer the research questions: What interaction factors of live streaming influence knowledge learning effects in knowledge-sharing contexts?

Knowledge-sharing Live Streaming

Live streaming has become a dominant real-time interactive digital media format worldwide, with diverse ecosystems across regions and platforms. Among its evolving forms, knowledge-sharing live streaming has emerged as a distinct and increasingly important genre. The global live streaming market continues to grow rapidly, projected to reach USD 76.86 billion in 2025, USD 97.39 billion in 2026, and USD 318.56 billion by 2031 (Mordor Intelligence 2026). Meanwhile, enterprises are increasingly adopting live streaming as a tool to support collaboration and knowledge sharing, further expanding its influence (360iResearch 2026).

Existing research has examined live streaming from perspectives such as user motivation, interaction mechanisms, engagement behaviors, and educational outcomes, spanning a wide range of content genres (see Appendix A). To contextualize knowledge-sharing live streaming, prior studies have explored knowledge sharing across organizational settings (Lukito and Perdhana 2023), educational domains, and social media environments (Zhao et al. 2020). With the rise of social media, knowledge sharing has shifted toward decentralized, user-generated environments, where knowledge is co-created through interaction, feedback, and networked communication. Within this context, knowledge-sharing live streaming has become a prominent mode of knowledge exchange, distinguished by real-time interaction (Yang et al. 2022).

Live streaming integrates scholarly content with mass appeal, lowers professional barriers, and promotes accessible learning, while requiring specific content and operational strategies(Tsinghua University 2024). Its interaction is characterized by affordances such as two-way communication, active control, synchronicity, and social presence (Chen and Liu 2024). Prior studies conceptualize interactivity as multidimensional, commonly including these elements, while social presence reflects users' perceived sense of warmth and togetherness in mediated environments (Hou et al. 2025).

Therefore, investigating knowledge-sharing live streaming is critical for understanding how these unique interactions shape knowledge dissemination and learning outcomes in contemporary digital environments.

Conceptual Background

Live Streaming Interaction

In recent years, live streaming has emerged as an emerging media form integrating communication, technology, and social relationships. Existing studies mainly examine e-commerce, entertainment, gaming, and virtual streamers, focusing on user motivations, interaction mechanisms, streamer traits, and platform systems. Guo et al. (2025) studied co-playing and stickiness, and Lv et al. (2022) compared virtual versus real streamers. Despite contextual differences across research, this line of inquiry ultimately aims to explain the determinants of successful live stream sessions, with a substantial body of question focusing on the interactive antecedents of live stream performance. First, interaction mechanisms such as two-way communication enhance trust and continuous intention, as shown through socio-technical models (Hou et al. 2024). Second, active control—users' ability to manage exchanged information—promotes trust, recommendation adoption, and interaction quality. Third, synchronicity, reflecting real-time information exchange, strengthens authenticity, emotional attachment, and user stickiness (Li et al. 2021). Fourth, social presence further promotes perceived diagnosticity, enjoyment, and social support, thereby fostering sustained user participation within live streaming contexts(Hou et al. 2025).

In summary, although studies discuss two-way communication, active control, synchronicity, and social presence (see Appendix B), little is known about learning outcomes in knowledge-sharing live streams, where motivations differ from monetary incentives. Thus, research on knowledge learning effects in knowledge-sharing live streaming remains scarce.

Knowledge Learning Effects

Existing research on knowledge-sharing live streaming remains limited, particularly regarding what enhances knowledge learning effects. Before further discussion, it is necessary to clarify the concept of knowledge learning effects. It is suggested that knowledge learning effects should not be conceived as a single dimensional factor but a more complicated factor. Their review indicates three relevant dimensions: knowledge acquisition, knowledge transfer, and knowledge sharing. Knowledge acquisition concerns learning and absorbing new knowledge (Akbar et al. 2024); knowledge transfer refers to internalizing external knowledge (De Wit-de Vries et al. 2019); and knowledge sharing focuses on re-creating and disseminating knowledge after internalization through behaviours such as forwarding, commenting, or re-posting (Fonseca and Barbosa 2021).

Recent studies examine knowledge learning effects in diverse scenarios(see Appendix B), including library services (Akbar et al. 2024), organizational management (Lukito and Perdhana 2023), and online communities (Jia 2025). Prior research generally focuses on mechanisms of acquisition, transfer, and sharing. For example, Akbar et al. (2024) found personality traits influence librarians' knowledge acquisition; Vries et al. (2019) highlighted trust and communication in university–enterprise knowledge transfer; and Jia (2025) revealed how communication visibility affects knowledge sharing quality in project teams.

Prior research views knowledge learning as multidimensional (acquisition, transfer, sharing). As live streaming grows, its interaction reshapes knowledge sharing; however, limited research has examined how its interactive affordances influence learning outcomes.

Hypothesis Development

Two-way Communication and Knowledge Learning Effects

Live streaming provides customers with a session where they can obtain real-time two-way interaction, where they can autonomously discuss what content to consume and how to obtain information, and they can get timely feedback. According to Hou et al. (2024), Two-way communication, plays an important role in promoting participation. In live streaming environments, two-way communication helps establish cognitive mutual trust between audiences and streamers, promoting audience recognition and absorption of knowledge, thereby improving knowledge learning effects. Therefore, we hypothesize:

H1: Two-way communication is positively correlated with knowledge learning effects.

Active Control and Knowledge Learning Effects

Active control emphasizes audiences' control over the experience of live streaming content (Hou et al. 2024). Existing research shows that active control can help learners receive information more selectively, thereby improving their attention and participation. In knowledge sharing live streaming, audiences' active control enables them to gain their learning experience based on their interests and needs, enhancing their perceived efficacy and learning motivation, making them more focused when receiving knowledge. Therefore, we hypothesize:

H2: Active control is positively correlated with knowledge learning effects.

Synchronicity and Knowledge Learning Effects

Synchronicity emphasizes the immediacy of information exchange, which can reduce cognitive interruption and maintain audience attention and interaction rhythm (Ou et al. 2014). In live streaming, synchronic information feedback enables audiences to quickly digest and understand live streaming

content, maintaining focus and absorption of knowledge. Therefore, synchronicity in knowledge sharing live streaming may significantly improve audiences' absorption and internalization of knowledge, thereby enhancing knowledge learning effects. Therefore, we hypothesize:

H3: Synchronicity is positively correlated with Knowledge Learning Effects.

Social Presence and Knowledge Learning Effects

Social presence refers to individuals' perception of the sense of existence in virtual environments and is widely studied in immersive media and human-computer interaction. Bai & Tan (2025) found that social presence strengthens emotional bonds between streamers and audiences. In knowledge-sharing live streaming, audiences seek both interpersonal connection and cognitive benefits. Higher social presence can facilitate deeper engagement in knowledge acquisition, increasing attention, motivation, and cognitive involvement, thereby strengthening knowledge learning effects. Therefore, we hypothesize:

H4: Social Presence is positively correlated with Knowledge Learning Effects.

Based on the above, the research framework is proposed (Figure 1).

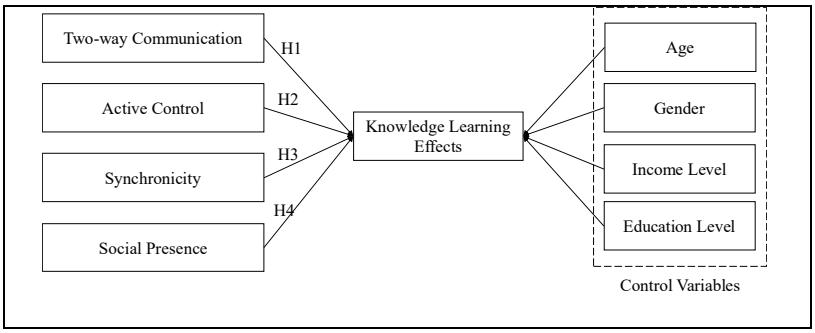


Figure 1. Research Framework

Research Methodology

This study employed a two-phase mixed-method approach. To test our proposed research framework, the first phase involved a quantitative investigation through survey questionnaire. Specifically, we conduct joint analysis using structural equation modelling (PLS-SEM), which can be used to verify the proposed research framework. The second phase is the qualitative investigation through focus group to provide in-depth insights to explain the influences examined in the first phase.

Questionnaire Design

This study used a questionnaire survey, with measurement items adapted from prior research for two-way communication, active control, and synchronicity (Hou et al. 2024; Ou et al. 2014); social presence (Cyr et al. 2009); and knowledge learning effects (Ma et al. 2023) (see Appendix C). Following J. N. Zhu et al. (2023), “In my estimate” was added to some items. Respondents rated their most recent knowledge-sharing live streaming experience on a 7-point Likert scale. Age, gender, income, and education level were included as controls, and the questionnaire was translated, expert-reviewed, and back-translated for accuracy.

Data Collection

The questionnaire was distributed via Credamo, a Chinese academic survey platform. A screening question ensured respondents had watched knowledge-sharing live streaming before; only those who answered “yes” continued . In total, we received 550 attempts of the survey, and 536 of them had the experience of watching knowledge sharing live streaming. After removing responses with missing values and those with the same answer to all questions, 526 valid responses were used for data analysis. The sample is relatively balanced in gender and predominantly aged 21–40, with a range of education levels and income groups represented (see Appendix D).

Data Analysis

Assessment of measurement model

We used SmartPLS4.0 to conduct SEM and assess measurement quality Internal consistency was assessed using Cronbach's alpha, rho_a, and rho_c. As shown in Appendix D , Cronbach's alpha for all constructs above 0.700 (0.714–0.800), confirming reliability (Ab Hamid et al. 2017). Convergent and discriminant validity were assessed, with all AVE values above 0.5 (0.513–0.660), confirming convergent validity (Fornell and Larcker 1981). Discriminant validity was supported as all HTMT values were below 0.85(0.023-0.797) (Ab Hamid et al. 2017) . VIF values were below 3 (1.262-1.776), showing no multicollinearity (Neter et al. 1990). To address common method bias, we (1) used established items, (2) ensured anonymity and no right/wrong answers, and (3) applied the CFA marker technique (Bowling 2010; Williams et al. 2010) with a four-item tardiness-punishment marker.

Assessment of structural model

Structural model results (Figure 2) show that two-way communication is non-significant (β = –0.019, P = 0.619; H1 not supported), while active control (β = 0.153, P = 0.002; H2), synchronicity (β = 0.207, P < 0.001; H3), and social presence (β = 0.387, P < 0.001; H4) have positive effects on learning. Detailed values are in the Appendix D. Control variables are non-significant, and the marker variable indicates minimal common method bias.

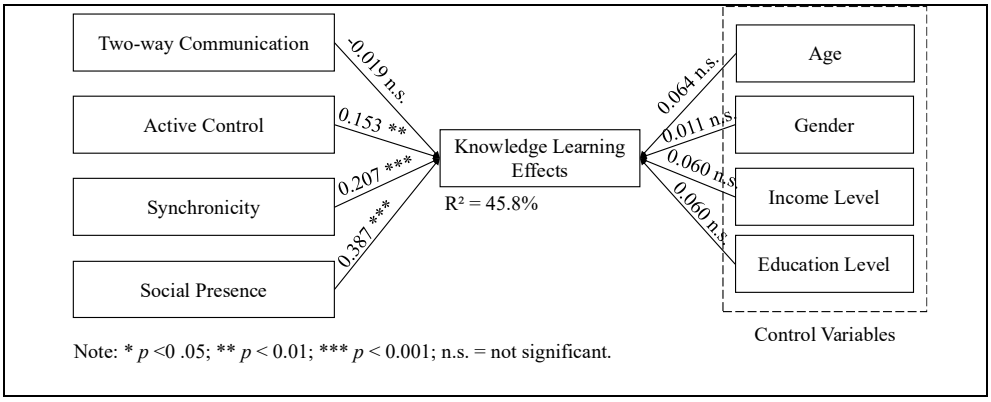


Figure 2. Research Model Data Results

Supplementary Qualitative Study

Building on the SEM results, social presence significantly positively predicted learning outcomes, whereas two-way communication exerted an insignificant effect. Notably, although two-way communication is generally regarded as a core dimension of interactivity and has been validated as effective in prior research, it failed to enhance knowledge learning effects in the knowledge sharing live streaming setting of this study. This intriguing inconsistency with existing literature motivated the present study to conduct semi-structured focus group interviews. We further explored participants' authentic viewing perceptions and experiences, so as to unpack the underlying reasons for such divergent findings and clarify the intrinsic mechanisms at play.

Participants and Data Collection

Participants were recruited via theoretical and snowball sampling, targeting experienced profiles and expanding by referral; twenty joined 30-minute guided discussions based on the questions about users' viewing motivations and interaction needs (detailed questions are shown in the Appendix E). All interviews were recorded, transcribed, and analyzed thematically.

Key Findings

Appendix E presents representative quotations illustrating the relationships between social presence, two-way communication, and knowledge learning effects.

Focus group results suggest that social presence enhances learning by increasing feelings of value, immersion, and emotional resonance through anchors' non-verbal cues, timely responses, professional performance, and visual aids, thereby improving engagement and attention. For two-way communication, its impact appears context-dependent, which may explain its non-significant overall effect. According to focus group results, it can support learning when interactions are content-focused, authentically delivered, or used in relaxed contexts (e.g., popular science), and even utilitarian interactions may prolong engagement; however, it may hinder learning when off-topic or in high-density, serious streams (e.g., exam preparation). Platform characteristics also matter, as users with clear learning goals may have low demand for non-content interaction. Overall, scenario matching and contextual heterogeneity likely shape its effectiveness.

Discussion

This study clarifies the differing influence of core situational factors on perceived knowledge learning effects. PLS-SEM results show that social presence has the strongest impact, followed by synchronicity and active control, with all three exerting significant positive effects on learning outcomes. This conclusion is consistent with the existing mainstream research findings. Prior research shows that social presence, synchronicity, and active control enhance cognitive processing in online and live streaming contexts. (Guan et al. 2022a; Luo et al. 2025; Yang et al. 2022). It is evident that perceived social presence, real-time information delivery, and learners' autonomous control exert stable and universally positive effects within knowledge-sharing live streaming contexts, consistent with findings across other streaming settings.

Different from the traditional view, the data of this study shows that two-way communication has no significant impact on knowledge learning effects (β=-0.019, p=0.619). Combined with the qualitative analysis of focus groups, the actual effectiveness of two-way communication has strict situational boundaries: be put forward that only in relaxed popular science live streams can in-depth and sincere interaction around core knowledge contribute to learning; in contrast, irrelevant chat interactions, rigid and formal exchanges, or redundant interactive behaviors in high-knowledge-density live streams such as exam preparation will distract the audience and interfere with in-depth knowledge learning. At the same time, in the platform environment where users have highly clear learning goals, the audience demand for non-knowledgeable interaction itself is relatively low, which further weakens the actual value of two-way communication. Notably, this finding deviates from conventional conclusions drawn in other live streaming contexts (e. g. entertainment and e-commerce), where two-way communication is widely confirmed to positively shape audience behavioral responses(Cheng et al. 2025; Guan et al. 2022b). By contrast, within knowledge-sharing live streaming scenarios, the effectiveness of two-way communication is conditional and requires deliberate, context-based regulation rather than universal application. This result clarifies the situational boundaries of two-way communication functionality, and compensates for the overgeneralized inferences of existing studies.

Implication

This study offers two main contributions. From a theoretical perspective, it extends live streaming research beyond purchase and entertainment contexts by focusing on knowledge learning effects and proposing a framework that explains how different dimensions of interactivity translate into learning effects in knowledge-sharing livestreams. It also refines the boundary conditions under which two-way communication is effective and highlights the differentiated roles of key interaction features. From a practical perspective, the findings suggest that platforms should strengthen social presence through clearer interaction displays and more intuitive interfaces, while improving latency, synchronisation, and system stability to support sustained attention. In addition, practitioners should clarify audiences' learning goals and design interactive content that closely aligns with core knowledge themes.

Limitations and Future Research

This study has several limitations that suggest directions for future research. First, the sample mainly consists of urban, highly educated users in China and is drawn from a single live streaming platform. Future research may extend this work by examining diverse knowledge-sharing contexts, such as knowledge-based e-commerce, and conducting multi-group analyses across platforms and content types to assess whether the identified interaction mechanisms, particularly the role of two-way communication, vary across settings. Second, as the questionnaire relies on self-reported perceptions, future research may adopt experimental designs and objective measures to better capture actual learning outcomes.

References

360iResearch. 2026. "Live Streaming Software Market Size & Share 2026-2032," *360iResearch*, , March 22.

Ab Hamid, M. R., Sami, W., and Mohmad Sidek, M. H. 2017. "Discriminant Validity Assessment: Use of Fornell & Larcker Criterion versus HTMT Criterion," *Journal of Physics: Conference Series* (890:1), IOP Publishing, p. 012163.

Akbar, A., Malik, A., and Warraich, N. F. 2024. "Personality Traits of Academic Librarians and Their Knowledge Acquisition and Application Behavior," *Libri* (74:3), pp. 305–321.

Bai, K., and Tan, K. H. 2025. "The Impact of Cognitive-Emotional Mechanisms on Online Presence and Livestreaming Satisfaction: A Two-Stage Survey Study," *Young Consumers: Insight and Ideas for Responsible Marketers*, pp. 1–21.

Bolun, Z., Yan, Z., and Minghui, J. 2025. "The Configurational Impact of E-Commerce Live Streaming Interactivity on Consumer Engagement Behavior," *Asia Pacific Journal of Marketing and Logistics* (37:3), pp. 631–649.

Bowling, N. A. 2010. "Effects of Job Satisfaction and Conscientiousness on Extra-Role Behaviors," *Journal of Business and Psychology* (25:1), pp. 119–130.

Chen, J., and Liu, H. 2024. "Modeling Interaction Behavior and Preference Decline for Live Stream Recommendation," *Decision Support Systems* (179), p. 114146.

Cheng, X., Xu, Y., Yang, B., and Liu, Y. 2025. "Understanding Streamers' Intention to Live Stream in the Context of Rural Live Streaming Commerce: A Mixed-Methods Study," *Internet Research* (36:2), pp. 822–846.

Cram, W. A., and Marabelli, M. 2018. "Have Your Cake and Eat It Too? Simultaneously Pursuing the Knowledge-Sharing Benefits of Agile and Traditional Development Approaches," *Information & Management* (55:3), pp. 322–339.

Cyr, D., Head, M., and Ivanov, A. 2009. "Perceived Interactivity Leading to E-Loyalty: Development of a Model for Cognitive–Affective User Responses," *International Journal of Human-Computer Studies* (67:10), pp. 850–869.

De Wit-de Vries, E., Dolfsma, W. A., Van Der Windt, H. J., and Gerkema, M. P. 2019. "Knowledge Transfer in University–Industry Research Partnerships: A Review," *The Journal of Technology Transfer* (44:4), pp. 1236–1255.

Fonseca, L., and Barbosa, S. 2021. "Knowledge Sharing Live Streams: Real-Time and On-Demand Engagement:," in *Proceedings of the 23rd International Conference on Enterprise Information Systems*, Online Streaming, --- Select a Country ---: SCITEPRESS - Science and Technology Publications, pp. 441–450.

Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), SAGE Publications Inc, pp. 39–50.

Gu, Y., Cheng, X., and Shen, J. 2023. "Design Shopping as an Experience: Exploring the Effect of the Live-Streaming Shopping Characteristics on Consumers' Participation Intention and Memorable Experience," *Information & Management* (60:5), p. 103810.

Guan, Z., Hou, F., Li, B., Phang, C. W., and Chong, A. Y.-L. 2022. "What Influences the Purchase of Virtual Gifts in Live Streaming in China? A Cultural Context-Sensitive Model," *Information Systems Journal* (32:3), pp. 653–689.

Guo, S., Chen, H., Dai, B., Wang, M., Liu, Shuo, Zhang, C., Liu, Shan, and Lu, X. 2025. "Influential Factors of Streamer Popularity in the Live Streaming Platform," *Journal of Retailing and Consumer Services* (84), p. 104194.

Hou, F., Guan, Z., Li, B., and Chong, A. Y. L. 2025. "Streaming Together: Unraveling Social Experiences in Social Live Streaming Services," *Information & Management* (62:8), p. 104230.

Hou, F., Li, B., Guan, Z., Chong, A. Y. L., and Phang, C. W. 2024. "The Role of Para-Social Relationship in Live Streaming Virtual Gift Purchase: A Two-Stage SEM-Neural Network Analysis," *Internet Research* (34:6), pp. 2298–2340.

Jia, J. 2025. "Antecedents of Knowledge Sharing in Online Collective Spaces: Combining Social Power Theory and Knowledge Governance Theory," *Information Technology & People*.

Li, Y., Li, X., and Cai, J. 2021. "How Attachment Affects User Stickiness on Live Streaming Platforms: A Socio-Technical Approach Perspective," *Journal of Retailing and Consumer Services* (60), p. 102478.

Lu, Z., Xia, H., Heo, S., and Wigdor, D. 2018. "You Watch, You Give, and You Engage: A Study of Live Streaming Practices in China," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal QC Canada: ACM, April 21, pp. 1–13.

Lukito, D., and Perdhana, M. S. 2023. "Investigating the Relationship of Change Leadership, Knowledge Acquisition, and Firm Performance in Digital Transformation Context," *Quality - Access to Success* (24:194).

Luo, X., Lim, W. M., Cheah, J.-H., Lim, X.-J., and Dwivedi, Y. K. 2025. "Live Streaming Commerce: A Review and Research Agenda," *Journal of Computer Information Systems* (65:3), Taylor & Francis, pp. 376–399.

Lv, J., Cao, C., Xu, Q., Ni, L., Shao, X., and Shi, Y. 2022. "How Live Streaming Interactions and Their Visual Stimuli Affect Users' Sustained Engagement Behaviour—A Comparative Experiment Using Live and Virtual Live Streaming," *Sustainability* (14:14), p. 8907.

Ma, S., Li, D., Wang, Y., and Han, M. S. 2023. "How Does IT Capability Affect Knowledge Acquisition in the Presence of a Dominant Supplier?," *Journal of Knowledge Management* (27:8), pp. 2090–2112.

Mordor Intelligence. 2026. "Live Streaming Market Size, Share Analysis, Forecast Report 2026 – 2031," *Mordor Intelligence*, , January 27.

Neter, J., Wasserman, W., and Kutner, M. H. 1990. "Polynomial R1egression," *Neter J, Wasserman W, Kutner MH. Applied Linear Statistical Models. Boston: Irwin*, pp. 315–41.

Ou, C. X., Pavlou, P. A., and Davison, R. M. 2014. "Swift Guanxi in Online Marketplaces: The Role of Computer-Mediated Communication Technologies," *MIS Quarterly* (38:1), Management Information Systems Research Center, University of Minnesota, pp. 209–230.

Tsinghua University. 2024. "清华学术报告：学无界，共此间-短视频直播与知识学习 _光明网."

Williams, L. J., Hartman, N., and Cavazotte, F. 2010. "Method Variance and Marker Variables: A Review and Comprehensive CFA Marker Technique," *Organizational Research Methods* (13:3), SAGE Publications Inc, pp. 477–514.

Yang, S., Yim, J., Kim, J., and Shin, H. V. 2022. "CatchLive: Real-Time Summarization of Live Streams with Stream Content and Interaction Data," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA: Association for Computing Machinery, April 29, pp. 1–20.

Zhang, J. (Kristen), and Liu, R. 2024. "Why Do Chinese People Consume Video Game Live Streaming on the Platform? An Exploratory Study Connecting Affordance-Based Gratifications, User Identification, and User Engagement," *Telematics and Informatics* (86), p. 102075.

Zhao, Y., Zhang, X., Wang, J., Zhang, K., and Ordóñez de Pablos, P. 2020. "How Do Features of Social Media Influence Knowledge Sharing? An Ambient Awareness Perspective," *Journal of Knowledge Management* (24:2), pp. 439–462.

Zhu, J. N., Lam, L. W., Liu, Y., and Jiang, N. 2023. "Performance Pressure and Employee Expediency: The Role of Moral Decoupling," *Journal of Business Ethics* (186:2), pp. 465–478.

Appendix

Detailed information is presented in Appendices A–E.

KITS·ICEC·KIISS
2026 통합 국제학술대회

W2 Advanced Recommender Systems & Marketing

Session Chair : Sunghan Ryu
(Univ. of Nottingham, Ningbo)



W2

A Text Mining-based Analysis of Consumers' Online Pharmaceutical Purchasing Preferences

Jiali Xue, Jinxin Yang (Anhui University of Science and Technology)

This study examines consumers' online pharmaceutical purchasing preferences using 10,254 user reviews collected from JD Pharmacy between January 2024 and September 2025. By integrating Latent Dirichlet Allocation (LDA) topic modeling and lexicon-based sentiment analysis, the study identifies key consumer concerns and sentiment patterns in online pharmaceutical consumption.The results reveal four dominant themes: drug safety and professional reliability, drug efficacy and treatment effectiveness, price sensitivity and purchasing convenience, and service quality and user experience. These themes form a dual structure of product-related evaluation and service-related experience, indicating that consumer decision-making is jointly influenced by risk perception and service experience under information asymmetry.Sentiment analysis shows that most reviews are positive, while negative sentiments mainly relate to safety concerns, efficacy mismatch, logistics delays, supply shortages, and service issues. The findings provide insights for improving service design and enhancing trust in online pharmaceutical platforms.

KEY WORDS

Online pharmaceutical purchasing, Text mining, LDA topic modeling, Sentiment analysis, Consumer behavior

A Text Mining-based Analysis of Consumers' Online Pharmaceutical Purchasing Preferences

Xue, J. L.[a] *and Yang, J. X.*[b]
[a] *School of Economics and Management, Anhui University of Science and Technology*
E-mail: 2024201589@aust.edu.cn
[b] *chool of Economics and Management, Anhui University of Science and Technology*
E-mail: *jinxiny@aust.edu.cn*

Abstract

This study examines consumers' online pharmaceutical purchasing preferences using 10,254 user reviews collected from JD Pharmacy between January 2024 and September 2025. By integrating Latent Dirichlet Allocation (LDA) topic modeling and lexicon-based sentiment analysis, the study identifies key consumer concerns and sentiment patterns in online pharmaceutical consumption.The results reveal four dominant themes: drug safety and professional reliability, drug efficacy and treatment effectiveness, price sensitivity and purchasing convenience, and service quality and user experience. These themes form a dual structure of product-related evaluation and service-related experience, indicating that consumer decision-making is jointly influenced by risk perception and service experience under information asymmetry.Sentiment analysis shows that most reviews are positive, while negative sentiments mainly relate to safety concerns, efficacy mismatch, logistics delays, supply shortages, and service issues. The findings provide insights for improving service design and enhancing trust in online pharmaceutical platforms.

Keywords: Online pharmaceutical purchasing; Text mining; LDA topic modeling; Sentiment analysis; Consumer behavior

Introduction

With the advancement of the Healthy China Strategy and the rapid development of "Internet + Healthcare," pharmaceutical e-commerce platforms have become an increasingly important channel for drug consumption. In particular, since the COVID-19 pandemic, platforms such as JD Health and Ali Health have experienced significant growth in online pharmaceutical transactions(Anderson et al. 2022; Hu 2024) . As the pandemic transitions into a normalized phase, consumer behavior has shifted from emergency-driven purchasing to routine and experience-oriented decision-making. Accordingly, understanding how consumer preferences evolve in this context has become an important research issue.

Existing studies have widely applied text mining techniques, particularly Latent Dirichlet Allocation (LDA) and sentiment analysis, to investigate consumer perceptions in online pharmaceutical markets. Prior research has identified key concerns such as logistics efficiency, drug efficacy, and customer service quality(Liu et al. 2020; Zhao and Huang 2024) . Although these studies provide valuable insights, two limitations remain.First, most existing research focuses on the COVID-19 emergency period, which may not accurately reflect consumer behavior in the post-pandemic normalization stage. As consumer decision-making becomes more stable and experience-driven, the applicability of earlier findings requires further validation. Second, previous studies mainly analyze overall review data without considering heterogeneity across different drug categories, limiting a fine-grained understanding of consumer preference structures.

From a theoretical perspective, consumer behavior in online pharmaceutical contexts can be better understood through the lens of risk perception theory(Slovic 1987) and information asymmetry theory(Akerlof 1970), as online drug purchasing involves uncertainty regarding product quality, efficacy, and service reliability. However, empirical evidence integrating topic modeling and sentiment analysis to capture such multidimensional consumer concerns remains limited.

To address these gaps, this study analyzes 10,254 online reviews collected from JD Pharmacy between January 2024 and September 2025. By integrating LDA topic modeling and sentiment analysis, this study aims to (1) identify key consumer concern dimensions, (2) examine category-level heterogeneity in consumer preferences, and (3) analyze sentiment distribution patterns across topics. The findings contribute to a deeper understanding of consumer decision-making in digital healthcare environments and provide implications for platform optimization and service improvement.

Data Sources and Preprocessing

The dataset used in this study was collected from user reviews on the JD Pharmacy platform, covering the period from January 2024 to September 2025. The data were crawled using a Python 3.10-based web scraping program, resulting in an initial collection of 11,935 reviews. These reviews cover 17 pharmaceutical products (some of which correspond to different specifications of the same drug type), including medications for chronic diseases, common illnesses, and health-related products. This ensures diversity across different drug categories and enhances the representativeness of the dataset.

In the data preprocessing stage, a systematic cleaning procedure was conducted to improve data quality and ensure the reliability of subsequent analysis. First, reviews with excessively short length (fewer than 10 characters), duplicate entries, and garbled or meaningless text were removed to reduce noise and improve data validity. This filtering criterion is consistent with commonly adopted practices in text mining studies for eliminating low-information content(Côté et al. 2024). After data cleaning, a total of 10,254 valid reviews were retained.Next, Chinese word segmentation was performed using the Jieba toolkit (version 0.42.1) to transform unstructured text into meaningful lexical units, thereby improving the structural representation of Chinese textual data. A customized stopword list, combined with commonly used Chinese stopword dictionaries, was applied to remove function words, high-frequency non-informative terms, and platform-independent words that do not contribute meaningful semantic information. This step helps reduce noise and improve the quality of topic modeling results.To further standardize the text data, preprocessing operations such as punctuation removal, special character filtering, and elimination of meaningless symbols were applied to ensure consistency in text format. The processed text was then converted into a Bag-of-Words representation, and a Term Frequency (TF) matrix was constructed to serve as the input for subsequent LDA topic modeling analysis.Finally, after the above data cleaning and preprocessing procedures, a structured textual dataset was obtained, which was used as the input for both topic modeling and sentiment analysis in the subsequent empirical analysis.

Consumer Review Topic Analysis

Construction of LDA Topic Model

To identify the key topics consumers focus on when purchasing pharmaceuticals online, this study used the LDA (Latent Dirichlet Allocation) topic model. LDA introduces a latent variable, "topic," to represent each review as a probabilistic mixture of topics, effectively identifying the main concerns of consumers(Blei et al. 2003). During text feature extraction, a Count Vectorizer was used to construct a term frequency matrix, reducing lexical similarity between different topics to ensure strong differentiation(Jelodar et al. 2019).

To ensure the accuracy and interpretability of the LDA model, perplexity (Perplexity) was used to select the optimal number of topics. The model was trained with a range of 2 to 10 topics, and the topic number with the lowest perplexity was chosen. The optimal number of topics was determined to be 4 (see Figure 1), balancing topic differentiation and data coverage.

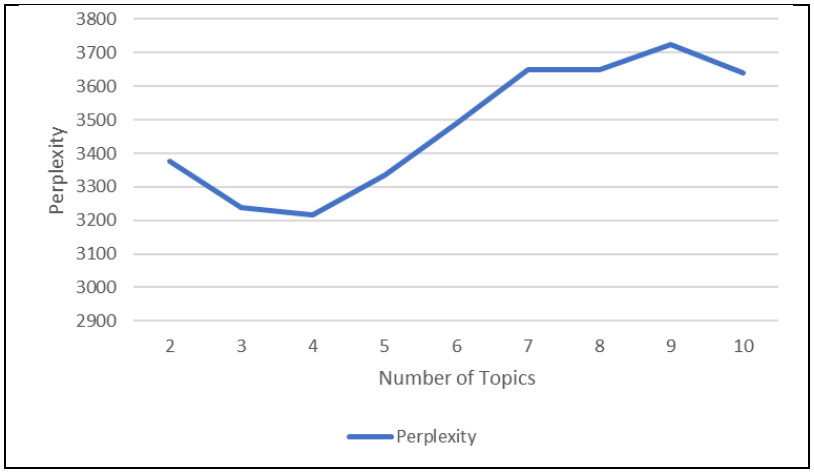


Figure 1. Selection of Optimal Number of Topics

Topic Identification and Result Analysis

The optimal number of topics identified in this study is 4, with the top 10 high-probability feature words extracted for each topic and summarized into themes (as shown in Table 1). The four main themes were identified as: drug professionalism and safety, drug efficacy, price and purchasing convenience, and service attitude and professionalism. From the summary of these four themes, it can be concluded that consumers' concerns can be categorized into two main dimensions: one pertains to the attributes of the product itself, including drug efficacy, safety, convenience, and the reliability of the purchasing channel; the other relates to the purchasing and service experience, including platform service quality, professionalism of the service, and the overall purchase experience. Overall, in the context of pandemic normalization, consumers' decision-making process for online pharmaceutical purchases still places a high value on drug efficacy and safety, while also raising higher expectations for channel security and service experience. Therefore, pharmaceutical e-commerce platforms need to not only ensure the quality and safety of drugs but also optimize their service processes to provide a better experience for consumers. This will help enhance consumer satisfaction and improve the platform's competitiveness.

Table 1. LDA Topic Model Results

	Top 10 High-Probability Feature Words	Topic Label
Topic 1	Parlin, Ada, Allergy, Blood Sugar, Cold Chain, Ointment, Tongrentang, Trouble, Patients, Original Drug	Drug Professionalism and Safety
Topic 2	Plaster, Constipation, Powder, Anti-inflammatory, Comprehensive, Independent, Dissolving, Healing, Stomach, Wound	Drug Efficacy
Topic 3	Baby,Headache,Fresh,Cold Medicine,Price Comparison, Backup, Platform, Physical Store, Prevention, Production Date	Price and Purchasing Convenience
Topic 4	Attitude, Excellent, Patient, Workmanship, Expectations, Professional, Thoughtful Service, Detailed, Perfect, Appearance	Service Attitude and Professionalism

Topic Distribution across Different Drug Categories

After categorizing and summarizing the online review topics, we analyzed the distribution of different drug categories across these themes to better understand consumers' preferences. The results in Table 2 show significant differences in consumers' focal points for various drug categories. For endocrine and nasal medications, Theme 1 is the most prevalent, indicating that consumers prioritize professionalism and reliability. In contrast, cold medications and vitamins are more focused on Theme 3, highlighting consumers' concern for price, purchasing channels, and convenience. For digestive, neurological, and rheumatic drugs, Theme 2 dominates, reflecting consumers' focus on efficacy and symptom relief. Throat and oral care products, however, are more centered on Theme 4, emphasizing the importance of platform delivery, packaging, and customer service. Overall, consumers of chronic disease medications prioritize professionalism and safety, those purchasing everyday medicines emphasize price and convenience, while consumers of functional or health-related products focus more on shopping experience and service quality.

Table 2. Topic Distribution across Different Drug Categories

Drug Category	Topic 1	Topic 2	Topic 3	Topic 4
Endocrine Medications	0.505	0.062	0.178	0.255
Oral Medications	0.279	0.098	0.266	0.357
Throat Care Medications	0.161	0.103	0.308	0.428
Cold Medications	0.103	0.158	0.585	0.155
Pain Relief Medications	0.210	0.135	0.431	0.224
Digestive System Medications	0.167	0.416	0.247	0.170
Skin Medications	0.393	0.203	0.204	0.200
Neurological Medications	0.209	0.370	0.181	0.240
Systemic Digestive Medications	0.065	0.675	0.176	0.084
Vitamins	0.092	0.147	0.522	0.239
Tonics and Supplements	0.337	0.155	0.233	0.275
Rheumatic and Trauma Medications	0.122	0.580	0.185	0.114
Nasal Medications	0.597	0.056	0.176	0.170

Sentiment Analysis of Consumer Online Reviews

Construction of Sentiment Analysis Model

The sentiment analysis method in this study is based on the Boson NLP Sentiment Lexicon, which covers a large number of commonly used terms in social networks and assigns each term a corresponding sentiment score to quantify its sentiment tendency. To improve the accuracy and specificity of the analysis results, this study also incorporates a dictionary of negation words and a dictionary of degree adverbs compiled from CNKI. In terms of assigning weights to degree adverbs, this study refers to the classification and scope definition of degree adverbs by Zhang Xiaoyan and Bai Yu in their research(Zhang and Bai 2022). The highest-level modifiers are assigned a value of 2.0, with other values being 1.7, 1.4, 1.1, 0.8, and 0.5 in descending order. The specific sentiment word weight distribution is shown in Table 3.

Table 3. Assignment of Degree Adverb Weights

Degree	Degree Adverb	Count (per unit)	Weight
Most	Extremely, highly, exceedingly, fully, absolutely, etc.	57	2.0
Over	Not too much, excessive, much, more than, etc.	28	1.7
Very	Especially, particularly, quite, very, truly, etc.	32	1.4
More	Still, relatively, especially, increasingly, etc.	35	1.1
Ish	Slightly, somewhat, a little, slightly more, etc.	29	0.8
Insufficently	Not much, barely, mildly, weakly, slightly, etc.	11	0.5

First, Chinese word segmentation was performed on each review, and sentiment words were matched to calculate sentiment scores for each word. Then, a weighted summation method was applied to calculate the overall sentiment score for each review.To improve classification accuracy, z-score normalization was applied to the sentiment scores, transforming them into data that follows a standard normal distribution. This ensured the robustness and reliability of the results. This method effectively quantifies consumers' sentiment toward online pharmaceutical purchasing and provides targeted recommendations for improvement.

Sentiment Analysis Results

To assess the accuracy of the sentiment analysis model constructed in this study, 100 random reviews were selected from the data sample as test samples for sentiment analysis. If the sentiment analysis result was positive, the review was categorized as positive; if the result was negative, the review was categorized as negative; and if the result was zero, the review was classified as neutral. Subsequently, each review's sentiment was manually labeled (neutral, positive, or negative). Finally, the sentiment analysis results were compared with the manually labeled data(Hartmann et al. 2023; Khoo and Johnkhan 2018). The results showed an accuracy of 83%, indicating that the model performs well in sentiment classification. Based on this, the model was applied to all cleaned review data, with the results shown in Table 4. The data reveals that positive reviews account for 88.3%, negative reviews account for 11.4%, and neutral reviews only account for 0.3%. This distribution suggests that, in the context of pandemic normalization, consumers generally hold a positive attitude toward online pharmaceutical purchasing, and online channels have gradually become a key choice for the public. By combining the results of LDA topic modeling and sentiment analysis, this study identified negative sentiment reviews

within each topic. Analyzing the content of these negative reviews and their corresponding topics helped summarize the issues consumers encounter when purchasing medications online.

Table 4. Number and Proportion of Different Sentiment Categories

Sentiment Category	Quantity	Proportion
Positive	9050	88.3%
Negative	1170	11.4%
Neutral	34	0.3%

Consumers face several issues when purchasing medications online. First, in terms of medication safety, due to a lack of adequate understanding of their own physical condition and medications, consumers may experience adverse reactions such as allergies. This issue is compounded by false advertising from sellers and insufficient services from online pharmacies. Platforms have not established effective allergy history screening mechanisms, and pharmacists do not provide sufficient warnings regarding contraindications for special groups. Second, in terms of drug efficacy, consumers sometimes find that the actual effects of cold medicines and topical drugs do not match their expectations, indicating a mismatch between the advertised and actual efficacy of some products. Third, in terms of logistics and delivery, consumers often experience delays in receiving urgent medications, which leads to missed treatment windows. Fourth, in terms of drug supply, especially during seasonal disease peaks, some medications, like cold medicines, may be out of stock. The platform has not implemented a dynamic inventory warning system. Fifth, in terms of service attitude and professionalism, the pharmacist consultation service on the platform has become a procedural formality, failing to provide personalized medication plans for complex conditions, which increases the risk of patients using the wrong medication. Finally, in terms of post-sale rights protection, the platform has not established proactive recall and compensation mechanisms, resulting in high costs for consumers to protect their rights. These issues indicate that there is significant room for improvement in providing safe, professional, and efficient services on pharmaceutical e-commerce platforms.

Conclusion and Implications

Research Conclusion

Based on 10,254 user reviews collected from JD Pharmacy between January 2024 and September 2025, this study employs Latent Dirichlet Allocation (LDA) topic modeling and sentiment analysis to systematically investigate consumer concerns, preference heterogeneity, and emotional patterns in online pharmaceutical purchasing under the post-pandemic normalization context.The results indicate that consumer reviews are primarily organized around four core dimensions, including drug safety and professional reliability, drug efficacy and treatment effectiveness, price sensitivity and purchasing convenience, and service quality and user experience. These dimensions further reveal an underlying structural logic of consumer evaluation, which can be summarized into two higher-order categories: product-related evaluation and service-related experience. Specifically, product-related evaluation reflects consumers' concerns regarding uncertainty in drug safety and therapeutic effectiveness, while service-related experience captures perceptions of pricing rationality, logistics efficiency, and platform service quality.From a sentiment perspective, the majority of reviews are positive, suggesting that consumers generally hold favorable attitudes toward online pharmaceutical platforms in the post-pandemic normalization stage. However, negative reviews are primarily concentrated on issues such as uncertainty in drug safety, mismatches between expected and actual drug efficacy, logistics delays, drug shortages, and insufficient service responsiveness. These findings suggest that negative consumer experiences are largely driven by perceived risk and expectation gaps under conditions of information

asymmetry.Overall, online pharmaceutical consumption is not solely driven by price or efficacy considerations, but rather represents a complex decision-making process shaped by the joint influence of risk perception and service experience in uncertain informational environments.

Theoretical Implications

This study contributes to the literature on digital healthcare consumer behavior in several ways. First, it extends information asymmetry theory to the context of online pharmaceutical consumption, demonstrating that consumers face inherent disadvantages in evaluating drug quality and effectiveness due to limited medical expertise, which increases reliance on platform credibility and professional assurance.Second, the study deepens the explanatory power of risk perception theory in digital health environments by showing that consumers simultaneously evaluate both product-related risks (e.g., drug safety and efficacy uncertainty) and service-related risks (e.g., logistics reliability and service responsiveness), which jointly shape their overall attitudes and decision-making behavior.Furthermore, by integrating topic modeling and sentiment analysis, this study proposes an analytical framework for extracting multidimensional consumer insights from large-scale unstructured textual data, thereby contributing to methodological advancements in text mining applications within healthcare and e-commerce research.

Managerial Implications

This study provides several practical implications for pharmaceutical e-commerce platforms.First, platforms should strengthen drug safety assurance and trust-building mechanisms. Given consumers' strong concerns regarding drug safety uncertainty, improving transparency of drug information and enhancing professional pharmacist consultation services are critical to reducing perceived uncertainty.Second, logistics systems and supply chain responsiveness should be further optimized. The prevalence of negative sentiment related to delivery delays and stock shortages highlights the need for dynamic inventory management and emergency logistics mechanisms to improve service reliability.Third, improving service professionalism and responsiveness is essential. The findings suggest that service quality significantly influences overall consumer experience, particularly in complex medication scenarios where professional guidance can reduce misuse risks and enhance trust.Finally, platforms should leverage user review data and behavioral analytics to improve personalized recommendation systems, thereby enhancing matching efficiency between consumer needs and pharmaceutical products and improving user satisfaction and platform loyalty.

Limitations and Future Research Directions

Despite its contributions, this study has several limitations. First, the dataset is limited to a single platform (JD Pharmacy), which may restrict the generalizability of the findings. Future research could incorporate multi-platform datasets to enhance external validity and enable cross-platform comparison.Second, the sentiment analysis approach is lexicon-based, which, although interpretable, has limitations in capturing contextual semantics, sarcasm, and implicit emotional expressions. Future studies could adopt deep learning or transformer-based models to improve sentiment classification accuracy.Finally, this study relies on cross-sectional review data and does not capture the dynamic evolution of consumer preferences over time. Future research could integrate longitudinal behavioral data to explore the temporal dynamics of consumer decision-making in online pharmaceutical environments.

References

Akerlof, G. A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* (84:3), pp. 488-500.

Anderson, E. G., Lopez, J., and Parker, G. G. 2022. "Leveraging Value Creation to Drive the Growth of B2b Platforms," *Production and Operations Management* (31:12), pp. 4501-4514.

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent Dirichlet Allocation," *Journal of machine Learning research* (3:1), pp. 993-1022.

Côté, P.-O., Nikanjam, A., Ahmed, N., Humeniuk, D., and Khomh, F. 2024. "Data Cleaning and Machine Learning: A Systematic Literature Review," *Automated Software Engineering* (31:2), pp. 54.

Hartmann, J., Heitmann, M., Siebert, C., and Schamp, C. 2023. "More Than a Feeling: Accuracy and Application of Sentiment Analysis," *International Journal of Research in Marketing* (40:1), pp. 75-87.

Hu, S. 2024. "Research on Influencing Factors of Pharmaceutical E-Commerce Sales Based on Web Crawler and Support Vector Machine," *Transactions on Computer Science and Intelligent Systems Research* (4:1), pp. 48-59.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., and Zhao, L. 2019. "Latent Dirichlet Allocation (Lda) and Topic Modeling: Models, Applications, a Survey," *Multimedia tools and applications* (78:11), pp. 15169-15211.

Khoo, C. S., and Johnkhan, S. B. 2018. "Lexicon-Based Sentiment Analysis: Comparative Evaluation of Six Sentiment Lexicons," *Journal of Information Science* (44:4), pp. 491-511.

Liu, J., Zhou, Y., Jiang, X., and Zhang, W. 2020. "Consumers' Satisfaction Factors Mining and Sentiment Analysis of B2c Online Pharmacy Reviews," *BMC Medical Informatics and Decision Making* (20:1), pp. 194.

Slovic. 1987. "Perception of Risk," *Science* (236:4799), pp. 280.

Zhang, S., and Bai, Y. 2022. "Chinese Online Comments Sentiment Analysis Based on Weighted Char-Word Mixture Word Representation," *Application Research of Computers* (39:1), pp. 31-36.

Zhao, X., and Huang, Z. 2024. "A Method for Exploring Consumer Satisfaction Factors Using Online Reviews: A Study on Anti-Cold Drugs," *Journal of Retailing and Consumer Services* (81:1), pp. 103895.

W2

When AI Replies to Complaints: How Differences in Large Language Models Shape Consumers' Perceptions of Fairness and Brand Trust.

Li Suo (Huizhou University),
Sijing Yang, Magnum Man-Lok Lam, Christina Wing-Yan Wong (The Hong Kong Polytechnic University)

This study investigates how differences across large language models shape responses to negative e-commerce reviews and how their distinct communication styles influence consumers' perceptions of fairness and subsequent brand trust. As artificial intelligence increasingly handles customer complaints, model-specific characteristics can significantly affect post-failure trust recovery, yet such cross-model differences remain underexplored. Grounded in justice theory, this research employs a two-stage mixed-method design: a content analysis of 100 authentic reviews and a planned experiment. Preliminary findings indicate that ChatGPT predominantly emphasizes tangible compensation, whereas DeepSeek focuses on expressing empathy and outlining clear resolution procedures. This research clarifies systematic differences between large language models and offers practical guidance for brands in selecting and deploying appropriate AI agents for service recovery.

KEY WORDS

AI service recovery, Justice theory, Brand trust, E-commerce

When AI Replies to Complaints: How Differences in Large Language Models Shape Consumers' Perceptions of Fairness and Brand Trust

Suo, L.[a], Yang, S.[b], Lam, M. M.-L.[c], and Wong, C. W.-Y.[d]
[a] Glorious Sun Guangdong School of Fashion, Huizhou University, Huizhou, China
E-mail: 603181613@qq.com
[b] School of Fashion and Textiles, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: sijing-alicia.yang@connect.polyu.hk
[c] School of Fashion and Textiles, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: magnumml.lam@polyu.edu.hk
[d] School of Fashion and Textiles, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: christina.wy.wong@polyu.edu.hk

Abstract

This study investigates how differences across large language models shape responses to negative e-commerce reviews and how their distinct communication styles influence consumers' perceptions of fairness and subsequent brand trust. As artificial intelligence increasingly handles customer complaints, model-specific characteristics can significantly affect post-failure trust recovery, yet such cross-model differences remain underexplored. Grounded in justice theory, this research employs a two-stage mixed-method design: a content analysis of 100 authentic reviews and a planned experiment. Preliminary findings indicate that ChatGPT predominantly emphasizes tangible compensation, whereas DeepSeek focuses on expressing empathy and outlining clear resolution procedures. This research clarifies systematic differences between large language models and offers practical guidance for brands in selecting and deploying appropriate AI agents for service recovery.

Keywords: AI service recovery, justice theory, brand trust, e-commerce

Introduction

In modern e-commerce, firms increasingly rely on large language models (LLMs) to handle customer complaints. This shift stems from the demand for scalable and cost-effective service delivery (Zhang et al., 2025). However, effective recovery remains challenging, as distinct LLMs generate systematically different response styles. When these styles misalign with consumer expectations, they can undermine perceptions of fairness and erode brand trust (Lee, 2025). Existing service recovery research has two primary limitations. First, most studies treat AI as a homogeneous category in comparisons with human agents, overlooking structural differences among LLMs, including variations in training data and architecture (Tao et al., 2024). Second, although justice theory has recently been applied to AI-driven service recovery, limited attention has been paid to how specific LLMs prioritize and convey its core dimensions (Agnihotri & Bhattacharya, 2024). As a result, brands lack empirical guidance on selecting LLMs that best facilitate trust recovery. Service recovery fundamentally seeks to restore consumers' perceptions of fairness following a service failure (Van Vaerenbergh & Orsingher, 2016). Justice theory thus offers a robust framework for examining how LLMs convey fairness cues (Blodgett et al., 1997; Tax et al., 1998). Consumers assess AI responses using the same three dimensions applied to human recovery: distributive justice (outcome fairness), procedural justice (process fairness), and interactional

justice (interpersonal fairness) (Blodgett et al., 1997; Lin et al., 2011). Different LLMs may emphasize distinct dimensions implicitly. For instance, a model optimized for problem-solving might highlight compensation, whereas one tuned for empathy could prioritize emotional validation (Agnihotri & Bhattacharya, 2024). To investigate this phenomenon empirically, we employ a two-phase mixed-methods design that addresses three research questions: (1) How do two leading LLMs systematically differ in their deployment of justice cues when responding to negative reviews? (2) How do these cross-model differences influence consumers' perceived fairness? (3) How does perceived fairness, in turn, affect brand trust?

Theoretical Framework

Justice theory has long served as a central lens in service recovery research, explaining how consumers evaluate fairness after a failure and how those evaluations shape satisfaction, trust, and future behavior (Tax et al., 1998). The theory distinguishes three dimensions. Distributive justice concerns the fairness of the recovery outcome. Procedural justice concerns the fairness of the processes used to resolve the complaint. Interactional justice concerns the fairness of interpersonal treatment (Blodgett et al., 1997; Lin et al., 2011). Brand trust is typically assessed using scales grounded in commitment-trust theory (Morgan & Hunt, 1994). Empirical studies have repeatedly shown that all three dimensions contribute positively to post-recovery satisfaction, with distributive justice exerting the strongest effect on repurchase intention, and interactional justice being especially potent in mitigating negative word-of-mouth (Lin et al., 2011). More recently, justice theory has been successfully applied to AI-driven service recovery. Emerging evidence indicates that traditional relationships may operate differently in chatbot settings; for example, empathetic expressions from bots can sometimes negatively interact with monetary compensation (Lee, 2025), and user connection depends heavily on chatbot type (Coker & Thakur, 2026).

Methodology

The study adopts a two-phase mixed-methods design. Phase 1 consists of a comparative content analysis of AI-generated service recovery responses, and Phase 2 (planned) examines their effects on consumer perceptions and brand trust. For Study 1, 100 authentic negative reviews were collected from a leading apparel brand across five major Chinese e-commerce platforms. Restricting the sample to a single brand controls for brand-level characteristics while preserving linguistic diversity across platforms (Chevalier & Mayzlin, 2006). All reviews were written in Chinese and had not received official responses at the time of collection. Each review was submitted once to ChatGPT-5 and once to DeepSeek-V3 using identical standardized prompts requesting a professional service recovery response on behalf of the brand. These two models were selected for their distinct design paradigms and wide accessibility in e-commerce. All responses were generated in Chinese under consistent prompting conditions. This procedure resulted in 100 responses per model and 200 responses in total. Coding followed established operationalizations of distributive, procedural, and interactional justice (Blodgett et al., 1997; Lin et al., 2011; Tax et al., 1998). Initial automated tagging using Qwen Plus, cross-validated with GPT-5, was used to identify candidate justice cues. A trained human coder then reviewed and finalized all coding decisions according to a predefined coding protocol to ensure conceptual consistency. Cross-model differences in justice cue deployment were evaluated using chi-square tests with effect size estimation. Study 2 will employ a between-subjects experimental design (N = 300) to test the causal effects of the two response styles on perceived fairness and brand trust. Perceived distributive, procedural, and interactional justice will be measured using established multi-item scales (Blodgett et al., 1997; Lin et al., 2011; Tax et al., 1998), and brand trust will be assessed using Morgan and Hunt's (1994) commitment-trust scale. This design allows clear examination of causal effects from different justice cue emphases.

Preliminary Results

Drawing on justice theory, we examine how the two models differentially emphasize distributive, procedural, and interactional fairness cues when responding to identical negative reviews. Distributive

justice cues appeared in 87% of ChatGPT responses compared with 68% of DeepSeek responses, suggesting a stronger outcome-oriented recovery emphasis. In contrast, procedural justice cues appeared in 92% of DeepSeek responses versus 76% of ChatGPT responses, indicating greater process transparency in DeepSeek's replies. The most pronounced divergence emerged in interactional justice, present in 98% of DeepSeek responses but only 52% of ChatGPT responses, reflecting substantial differences in interpersonal engagement. Chi-square analyses confirm that these cross-model differences are statistically significant across all three dimensions. ChatGPT more frequently frames service recovery through distributive restoration (e.g., immediate refunds and compensation), while DeepSeek emphasizes procedural clarity and interactional empathy. Building on the preliminary findings from Study 1, Study 2 will employ a controlled between-subjects experiment to rigorously test the causal pathways from LLM-specific justice cues through perceived fairness to brand trust.

Conclusion

Theoretically, this study advances the service recovery literature by moving beyond traditional justice theory applications and revealing how the inherent communication styles of different large language models shape consumers' perceptions of fairness, empathy, and overall recovery effectiveness. It contributes novel theoretical insights into algorithmic fairness, the design of AI service agents, and the critical role of model architecture in influencing customer evaluations during service encounters. Practically, the results give e-commerce brands concrete guidance for training and selecting AI service agents: they can emphasize distributive justice for transactional recovery, or strengthen interactional and procedural justice for deeper customer relationships. Limitations include the single-brand sample and the dichotomous coding scheme that does not capture cue intensity; future research should address these.

References

Agnihotri, A., & Bhattacharya, S. (2024). Chatbots' effectiveness in service recovery. *International Journal of Information Management, 76*, Article 102679. https://doi.org/10.1016/j.ijinfomgt.2023.102679

Blodgett, J. G., Hill, D. J., & Tax, S. S. (1997). The effects of distributive, procedural, and interactional justice on postcomplaint behavior. *Journal of Retailing, 73*(2), 185–210.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research, 43(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Coker, K. K., & Thakur, R. (2026). Not all chatbots connect equally: The moderating role of chatbot type. *Electronic Markets, 36*(13). https://doi.org/10.1007/s12525-026-00873-4

Lee, J. (2025). When bots' empathic expressions backfire: Exploring the negative moderation between chatbot empathic expressions and monetary compensation for angry customers. *Electronic Markets, 35*(65). https://doi.org/10.1007/s12525-025-00814-7

Lin, H.-H., Wang, Y.-S., & Chang, L.-K. (2011). Consumer responses to online retailer's service recovery after a service failure: A perspective of justice theory. *Managing Service Quality, 21*(5), 511–534. https://doi.org/10.1108/09604521111159807

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing, 58*(3), 20–38.

Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2024). Cultural bias and cultural alignment of large language models. *PNAS Nexus, 3*(9), Article pgae346. https://doi.org/10.1093/pnasnexus/pgae346

Tax, S. S., Brown, S. W., & Chandrashekaran, M. (1998). Customer evaluations of service complaint experiences: Implications for relationship marketing. *Journal of Marketing, 62*(2), 60–76.

Van Vaerenbergh, Y., & Orsingher, C. (2016). Service recovery: An integrative framework and research agenda. *Academy of Management Perspectives, 30*(3), 328–346. https://doi.org/10.5465/amp.2014.0143

Zhang, Y., Ding, Z., Sun, J., Zhao, X., Hu, X., & Yang, Z. (2025). Optimizing AI strategies in e-commerce customer service: An agent-based simulation. *Electronic Markets, 35*(77). https://doi.org/10.1007/s12525-025-00821-8

W2

Warmth over Realism: Why Affective Likeability Trumps Technical Anthropomorphism and Expertise in Virtual Live-Streaming.

Tien-Dat Bui, Chih-Hsiung Hu (Department of Industrial Engineering, National Formosa University, Taiwan),
Tien-Dat Nguyen (Department of Smart Industry Technology Research and Development,
National Formosa University, Taiwan),
Chih-Chin Liang (Department of Business Administration, National Formosa University, Taiwan).

This research investigates the “Warmth over Realism” phenomenon in virtual live-streaming within Vietnam’s high-context emerging market. Grounded in the Stimulus-Organism-Response (S-O-R) framework, the study examines how Likeability, Anthropomorphism, and Expertise influence Parasocial Relationships (PSR), Consumer Trust, and Impulse Buying Behavior (IBB). Analysis of 375 valid responses using PLS-SEM reveals that Likeability is the dominant driver of emotional bonding and trust. Paradoxically, technical attributes-Anthropomorphism and Expertise-failed to significantly enhance consumer trust, a result attributed to the “Uncanny Valley” effect and a perceived lack of social warmth in AI. These findings confirm that impulsive consumption in this landscape is fundamentally “entertainment-first”. Strategic implications suggest that brands should prioritize stylized, likeable AI personas over hyper-realistic designs to optimize engagement and sales in the domestic market.

KEY WORDS

Information Classification, Cost-Sensitive Learning, Asymmetric Risk, Misclassification Cost

W2

From Expansive to Suppressive Escape: Emotional Exhaustion and User Disengagement on Immersive Digital Platforms.

Peng He (Kookmin Univ.) and Jing Zhang (Kookmin Univ.)

This study examines how different forms of immersive escapism shape user disengagement on immersive digital platforms. We distinguish between expansive escapism and suppressive escapism, and propose emotional exhaustion as the key mechanism linking escapism to downstream withdrawal.

Using user-generated review data from an immersive social platform, we construct three continuous text-based variables-expansive escapism, suppressive escapism, and emotional exhaustion-through theory-informed human annotation and LLM-assisted scoring. We further construct multiple post-review behavioral indicators, including incremental playtime (⊿PT), relative continued-use ratio (R), and follow-up time (Fi), and use them to identify discontinued use through a rule-based classification. We expect suppressive escapism to be more strongly associated with emotional exhaustion and, in turn, with a higher likelihood of discontinued use, whereas expansive escapism should show a weaker or opposite pattern. Follow-up experiments will further test the causal relationships and boundary conditions suggested by the review-based evidence.

KEY WORDS

Immersive Digital Platforms, Emotional Exhaustion, Expansive Escapism, Suppressive Escapism

W2

From Online to Offline Visit: Exploring Cognitive Fit Effect of AI-Streamer and Background.

Weiyu Dai, Xusen Cheng, Zhangtong Gu, Yijun Yan, Lu Gao (Renmin University of China)

While AI-streamers are transforming live commerce, the congruency between virtual environments and physical service contexts remains under-researched. Grounded in cognitive fit theory, we employ a mixed-methods approach to investigate how aligning virtual streamers and stream backgrounds with service contexts enhances multi-dimensional presence to drive physical store visit intention. The study will provide a novel framework for AI-driven live commerce, effectively bridging virtual engagement with physical store traffic.

KEY WORDS

Live commerce, AI-streamer, presence, cognitive fit, offline visit intention

From Online to Offline Visit: Exploring Cognitive Fit Effect of AI-Streamer and Background

Dai, Weiyu[a], Cheng, Xusen[b], Gu, Zhangtong[c], Yan, Yijun[d] and Gao, Lu[e]
[a] School of Information, Renmin University of China
E-mail: weiyu.dai@ruc.edu.cn
[b] School of Information, Renmin University of China
E-mail: xusen.cheng@ruc.edu.cn
[c] School of Information, Renmin University of China
E-mail: 2023202299@ruc.edu.cn
[d]School of Information, Renmin University of China
E-mail: yanyijun@ruc.edu.cn
[e] School of Information, Renmin University of China
E-mail: gaolu216@163.com

Abstract

While AI-streamers are transforming live commerce, the congruency between virtual environments and physical service contexts remains under-researched. Grounded in cognitive fit theory, we employ a mixed-methods approach to investigate how aligning virtual streamers and stream backgrounds with service contexts enhances multi-dimensional presence to drive physical store visit intention. The study will provide a novel framework for AI-driven live commerce, effectively bridging virtual engagement with physical store traffic.

Keywords: Live commerce, AI-streamer, presence, cognitive fit, offline visit intention

Introduction

As AI-streamers increasingly replace traditional hosts in live commerce, they play a crucial role in driving offline traffic for services that require physical engagement. However, how to configure the live streaming environment, specifically, how to design the optimal AI-streamers' appearance (realistic or stylized) and how to choose the matching background, remain under explored. Emerging studies challenge the assumption that maximum anthropomorphic realism is always superior (Seymour et al. 2018, 2021). Rather, from the perspective of cognitive fit theory, service providers can only maximize their strategic returns by ensuring a congruency between the live-streaming visual cues and the underlying service context.

Literature Review and Research Model

Current literature mainly examines the standalone attributes of AI-streamers, leaving the interplay between virtual configurations and physical service contexts under-explored. This study addresses this gap by theorizing that the effectiveness of live commerce hinges on the cognitive fit among AI-streamers, AI-generated backgrounds, and service contexts. We argue that such congruency evokes consumers' heightened multi-dimensional presence that fosters offline physical store visit intention.

Cognitive Fit Theory

Cognitive fit theory suggests that problem-solving performance improves when information presentation aligns with task demands. The user's mental processing is streamlined when the presentation format precisely aligns with the task requirements, resulting in noticeably quicker and more accurate decision-making (Martínez-Molés et al. 2024).

AI-Streamers and AI-Generated Backgrounds

Going beyond the traditional focus on how product types shape live commerce strategies (Luo et al. 2025), this study highlights that offline service contexts play an equally vital role in optimizing the configuration of AI-streamers and AI-generated backgrounds. Offline services are categorized as utilitarian or experiential services based on their dominant value orientation. Utilitarian services prioritize functional tasks (e.g., electronic product repair services), while experiential services emphasize sensory or emotional engagement (e.g., theme park experiences).

To effectively represent these distinct service types in live commerce, AI-streamers and AI-generated backgrounds vary in visual fidelity. AI-streamers range from realistic visuals (Cheng et al. 2025; Teng et al. 2023), which foster trust through anthropomorphism, to stylized visuals that evoke imagination. AI-generated backgrounds allow offline providers to bypass the high costs of physical studio construction, spanning from physical store simulations to abstract environments. Based on cognitive fit theory, these visual cues' efficacy depends on their alignment with the service context, as incongruency elevates cognitive load and diminishes presence (Diao et al. 2025).

Presence

Live streaming configurations effectively boost consumer presence, which reduces psychological distance and uncertainty by providing high-fidelity cues that simulate physical reality (Yoganathan et al. 2021). Presence manifests across multiple dimensions, including social co-existence (Teng et al. 2023), spatial immersion in remote environments (Khenak et al. 2020), cognitive information processing (Martínez-Molés et al. 2024), and self-presence, which reflects an individual's perceived existence. Unlike prior research focusing on maximizing individual dimensions, we argue that driving offline visit intention depends on a balance between these facets.

Hypotheses

Therefore, we propose the following hypotheses:

H1: For utilitarian in-store service contexts, a realistic AI-streamer combined with a physical AI-generated background results in higher multi-dimensional presence (social, spatial, self, cognitive) than other combinations.

H2: For experiential in-store service contexts, a stylized AI-streamer combined with an abstract AI-generated background results in higher multi-dimensional presence than other combinations.

H3: Multi-dimensional presence has a positive effect on offline visit intention.

H4: The effect of multi-dimensional presence on offline visit intention depends on the service context.

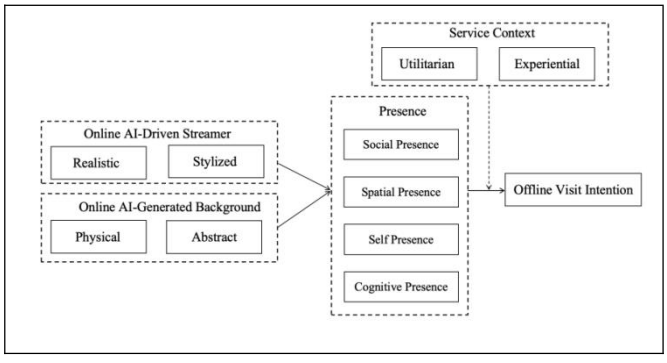


Figure 1. Research Model

Research Design

In the first phase, we developed an interview and recruited 30 participants experienced in AI-driven live streaming to validate our constructs and develop the measurement scale. Sessions were audio-

recorded and transcribed with consent. As no new themes emerged in the final interviews, we determined that theoretical saturation had been achieved.

In the second phase, hypotheses will be tested via a 2 × 2 × 2 between-subjects laboratory experiment. Participants will be randomly assigned to one of eight live-stream scenarios, followed by a survey measuring their perceived multi-dimensional presence and offline visit intention. Data will be analyzed by PLS-SEM to assess the moderated mediation effects. Control variables include age, gender, AI familiarity, live streaming viewing frequency, mental imagery, and cognitive style. Cognitive style is included as holists and analytic thinkers perceive contextual backgrounds differently, while age emerged from our interviews as a significant predictor of viewership and offline visit intention. Mental imagery is controlled to rule out the confounding effect of individual visualization differences on perceived presence (Bogicevic et al. 2019; Vazquez et al. 2023).

Conclusion

This research highlights that driving offline visit intention through live commerce depends on the fit between AI-streamers, AI-generated backgrounds, and offline service contexts. By applying cognitive fit theory to multi-dimensional presence (Gao et al. 2023; Khenak et al. 2020), we provide a novel framework for understanding how visual cues in live commerce bridge the virtual-physical divide. The findings will offer practical guidance for service providers to optimize AI-driven experiences that reduce consumer uncertainty and foster authentic connections aligned with their specific service context. Ultimately, this work will advance the discourse on human-AI interaction and the future of integrated online-offline service marketing.

Acknowledgements

Xusen Cheng is the corresponding author. We would like to thank the National Natural Science Foundation of China (Grant No. 72271236, 72434006) and Key Projects of Philosophy and Social Sciences Research of Chinese Ministry of Education (Grant No.23JZD022) for providing funding for this research.

References

Bogicevic, V., Seo, S., Kandampully, J., Liu, S. Q., and Rudd, N. 2019. "Virtual Reality Presence as a Preamble of Tourism Experience: The Role of Mental Imagery," *Tourism Management* (74).

Cheng, X., Zhang, S., and Mou, J. 2025. "Are You Caught in the Dilemma of Metaverse Avatars? The Impact of Individuals' Congruity Perceptions on Paradoxical Emotions and Actual Behaviors," *Decision Support Systems* (189), p. 114387.

Diao, Y., Yang, Q., Ge, S., Wang, N., and Lu, J. 2025. "A Study on the Role of Information Cues in E-Commerce Live Streaming: Evidence from Self-Reported and fNIRS Data," *Electronic Commerce Research* (25:4), pp. 3011–3045.

Gao, W., Jiang, N., and Guo, Q. 2023. "How Do Virtual Streamers Affect Purchase Intention in the Live Streaming Context? A Presence Perspective," *Journal of Retailing and Consumer Services* (73), p. 103356.

Khenak, N., Vézien, J., and Bourdot, P. 2020. "Spatial Presence, Performance, and Behavior between Real, Remote, and Virtual Immersive Environments," *IEEE Transactions on Visualization and Computer Graphics* (26:12), pp. 3467–3478.

Luo, L., Wang, Y., Zheng, Y., and Yuan, Y. 2025. "What to Say and How to Say It: Impacts of Content Type and Linguistic Style on Sales Performance on Live Streaming Platforms," *International Journal of Electronic Commerce* (29:2), pp. 235–271.

Martínez-Molés, V., Pérez-Cabañero, C., and Cervera-Taulet, A. 2024. "Examining Presence in Immersive Virtual Reality and Website Interfaces through the Cognitive Fit and Cognitive Load Theories," *International Journal of Contemporary Hospitality Management* (36:11), pp. 3930–3949.

Seymour, M., Riemer, K., and Kay, J. 2018. "Actors, Avatars and Agents: Potentials and Implications of Natural Face Technology for the Creation of Realistic Visual Presence," *Journal of the Association for Information Systems*, pp. 953–981.

Seymour, M., Yuan, L., Dennis, A., and Riemer, K. 2021. "Have We Crossed the Uncanny Valley? Understanding Affinity, Trustworthiness, and Preference for Realistic Digital Humans in Immersive Environments," *Journal of the Association for Information Systems* (22:3).

Teng, C.-I., Dennis, A. R., and Dennis, A. S. 2023. "Avatar-Mediated Communication and Social Identification," *Journal of Management Information Systems* (40:4), pp. 1171–1201.

Vazquez, E. E., Patel, C., Alvidrez, S., and Siliceo, L. 2023. "Images, Reviews, and Purchase Intention on Social Commerce: The Role of Mental Imagery Vividness, Cognitive and Affective Social Presence," *Journal of Retailing and Consumer Services* (74), p. 103415.

Yoganathan, V., Osburg, V.-S., H. Kunz, W., and Toporowski, W. 2021. "Check-in at the Robo-Desk: Effects of Automated Social Presence on Social Cognition and Service Implications," *Tourism Management* (85), p. 104309.

W2

MORE: A Multi-Agent Approach for Multi-Objective E-commerce Recommendation.

Yuxiang Pang, Di Wang (Xi'an Jiaotong University)

Recommender systems are essential to e-commerce because they support product discovery and purchase decisions. Recently, large language models (LLMs) have been incorporated into multiple stages of the recommendation pipeline to improve recommendation quality, owing to their strong natural language understanding and rich world knowledge. However, most existing methods still focus primarily on user intent, without fully accounting for the multiple objectives inherent in e-commerce recommendation. To bridge this gap, we propose MORE, an LLM-based multi-agent framework for e-commerce recommendation. MORE employs three specialized agents to score candidate items from the user, platform, and supply-side perspectives, and then aggregates their scores for reranking. We validate MORE across multiple LLM backends on Amazon Reviews'23 datasets. Experimental results show that, compared with conventional and single-agent recommenders, MORE achieves a better balance among user utility, platform-side value, and supply-side fairness.

KEY WORDS

Recommendation System, Large Language Model, Multi-Agent System

MORE: A Multi-Agent Approach for Multi-Objective E-commerce Recommendation

Yuxiang Pang[a] and Di Wang[a]
[a] School of Management, Xi'an Jiaotong University
E-mail: pollina99@stu.xjtu.edu.cn, wang.di@xjtu.edu.cn

Abstract

Recommender systems are essential to e-commerce because they support product discovery and purchase decisions. Recently, large language models (LLMs) have been incorporated into multiple stages of the recommendation pipeline to improve recommendation quality, owing to their strong natural language understanding and rich world knowledge. However, most existing methods still focus primarily on user intent, without fully accounting for the multiple objectives inherent in e-commerce recommendation. To bridge this gap, we propose MORE, an LLM-based multi-agent framework for e-commerce recommendation. MORE employs three specialized agents to score candidate items from the user, platform, and supply-side perspectives, and then aggregates their scores for reranking. We validate MORE across multiple LLM backends on Amazon Reviews'23 datasets. Experimental results show that, compared with conventional and single-agent recommenders, MORE achieves a better balance among user utility, platform-side value, and supply-side fairness.

Keywords: Recommendation System, Large Language Model, Multi-Agent System

Introduction

Recommender systems are central to e-commerce because they support product discovery and purchase decisions in large assortments (Xiao, 2007). While conventional recommender models have made substantial progress, they are typically built on discrete ID features and behavioral signals, which limit their ability to exploit open-world knowledge and to fully capture item semantics and users' underlying intents. In contrast, large language models (LLMs) provide rich semantic knowledge, stronger natural language understanding, and emergent reasoning capabilities from large-scale pretraining (Lin, 2025). These properties make them a promising foundation for enhancing recommendation beyond traditional interaction-based matching.

Existing LLM-based recommendation research has incorporated LLMs into various stages of the recommendation pipeline, such as feature engineering, feature encoding, scoring and ranking, interactive recommendation, and explanation generation (Jia, 2025; Chen, 2024). Yet broader architectural involvement does not necessarily imply broader optimization objectives. In practice, most of these methods still center on modeling user interests and improving relevance. This focus is incomplete for e-commerce recommendation, which is inherently multi-objective: recommendation decisions shape not only user utility, but also platform-side value and supply-side fairness (Jannach, 2023; Jiang, 2024).

Accordingly, e-commerce recommendation faces an important challenge: heterogeneous objectives are difficult to capture with a single undifferentiated preference signal. This raises a fundamental research question: *how can multiple objectives be jointly modeled and balanced within a unified recommendation framework?*

To address this problem, we propose **MORE**, an LLM-based **m**ulti-agent framework for multi-**o**bjective **r**ecommendation in **e**-commerce. As shown in Figure 1, MORE first uses a conventional retrieval module to construct a candidate set, and then employs multiple specialized LLM agents to

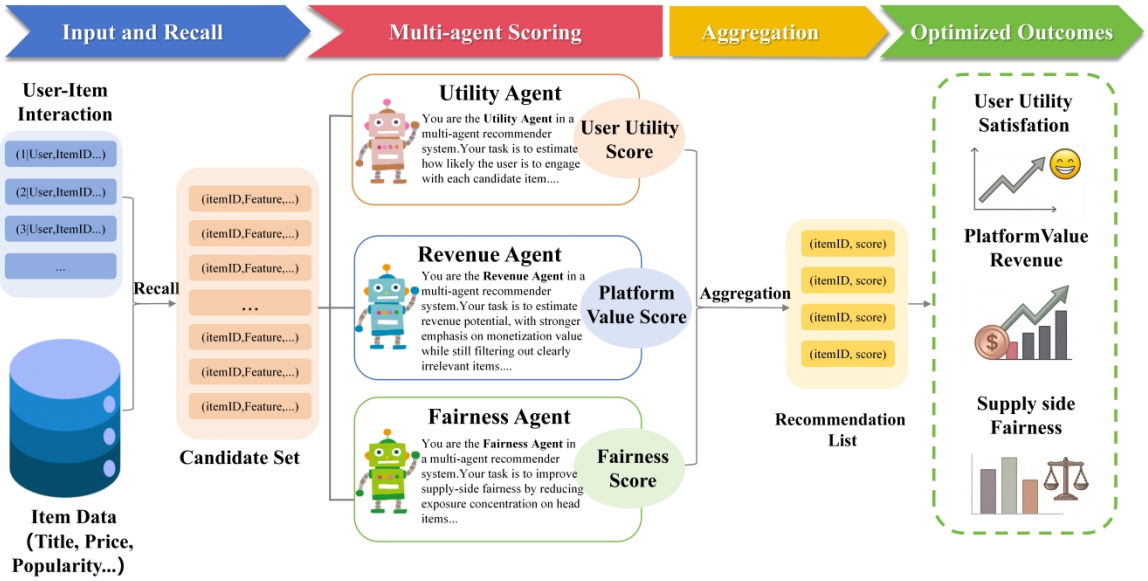


Figure 1. The Framework of MORE

score items from different objective-specific perspectives. The resulting scores are finally aggregated to produce a ranking that balances user utility, platform-side value, and supply-side fairness. Experimental results demonstrate that MORE consistently outperforms both traditional recommendation models and single-agent LLM-based methods across all three objectives.

In summary, our key contributions are as follows:

- We provide a formal characterization of the three core objectives in e-commerce recommendation, namely user utility, platform-side value, and supply-side fairness.
- We propose MORE, an LLM-based multi-agent framework that performs objective-specific item scoring and aggregation for e-commerce recommendation.
- We conduct extensive experiments showing that MORE consistently outperforms both traditional recommendation models and single-agent LLM-based methods across all three objectives.

Related work

E-commerce Recommendation

E-commerce recommendation has been widely adopted to help users discover relevant information and products in an increasingly overloaded information environment (Schafer, 2001). It spans multiple recommendation scenarios, including feed recommendation (Wu, 2021), item recommendation (Sharma, 2023), search recommendation (Si, 2023), and so on. Although these scenarios differ in form and function, they all involve balancing multiple objectives, including user utility, platform-side value, and supply-side fairness. The conflicts among these objectives make single-objective recommendation models inadequate and motivate the study of multi-objective and multi-agent methods.

LLM-based Recommendation

With the rise of LLMs, their rich open-world knowledge, strong semantic understanding, and reasoning capabilities have been increasingly explored in recommender systems (Zhao, 2024). Some studies leverage LLMs for feature engineering and feature encoding, using their semantic and world knowledge to construct richer and more accurate recommendation signals (Yin, 2023). Other works exploit their reasoning ability to score and rank candidate items, aiming to improve recommendation accuracy (Qu, 2016; Geng, 2022). In addition, LLMs have also enabled new paradigms such as conversational recommendation, which enriches user interaction and recommendation experience

(Zhang, 2023). However, most existing LLM-based recommendation methods mainly focus on optimizing user utility. Limited attention has been paid to jointly modeling the multiple objectives that arise in real-world recommendation, particularly the trade-off among user utility, platform-side value, and supply-side fairness.

Methodology

We propose MORE, an LLM-based multi-agent framework for multi-objective reranking in e-commerce recommendation. MORE operates on the recalled candidate set and performs list-level reranking through multiple agents, each corresponding to a specific recommendation objective. The framework includes two stages, including candidate construction via multi-channel recall and objective-specific scoring by LLM-based agents. The remainder of this section details these stages.

Let $u \in U$ denote a user, and let I denote the item set. For each user u, we use $H_u=[i_1,i_2,\ldots,i_T]$ to denote the historical interaction sequence, where $i_t \in I$ is an interacted item. Given H_u, the recall stage first constructs a candidate set $C_u \subset I$ and the reranking stage then outputs a reordered recommendation list π_u over C_u.

Recall

To construct a candidate set that supports subsequent multi-objective reranking, we adopt a multi-channel recall strategy that combines item-to-item (I2I) recall with random long-tail recall.

Item-to-item recall

Item-to-item recall is one of the most widely used retrieval strategies in industrial recommender systems. It retrieves items that are behaviorally or semantically similar to those in the user's historical interaction sequence, thereby producing a candidate set with strong preference relevance.

Formally, let $sim(i, j)$ denote the similarity between item i and item j. For user u, the I2I recall score s of a candidate item j is defined as $s(u, j)=\sum_{i \in H_u} sim(i, j)$, where the score accumulates the similarity between j and the items in the user history. Based on this score, the I2I recall set is constructed as $C_1=\mathrm{TopK_j}\ s(u, j)$, where C_1 represents the set of K candidate items with the largest values of $s(u, j)$. The key advantage of I2I recall is that it offers a strong relevance prior for downstream reranking. Since our goal is not to replace the retrieval stage but rather to study multi-objective optimization at the reranking stage, I2I recall provides an appropriate and practically meaningful retrieval backbone.

Random Long-Tail Recall

To give the reranker room to promote tail items, we introduce a random long-tail recall branch alongside I2I recall. Specifically, we define long-tail items as the bottom 90% of items ranked by interaction frequency, and randomly sample a subset of them to form an additional candidate set C_2. Unlike I2I recall, this branch is not designed to maximize immediate relevance, but to inject under-exposed items into the candidate pool. The final candidate set is obtained by merging the two recall sets.: $C_u=C_1 \cup C_2$.

Multi Agent Reranking Framework

E-commerce reranking is inherently multi-objective, as it must balance user preference, platform value, and supply-side fairness. Modeling these heterogeneous objectives with a single scoring function is difficult and often leads to implicit and uncontrollable trade-offs. To address this, we adopt a multi-agent design that decomposes reranking into multiple objective-specific scoring processes, each responsible for one aspect of the decision.

Shared Structured Context

To ensure that differences in model outputs arise from objective design rather than input variation, and to make the scores produced by different agents directly comparable, MORE keeps all parts of the input identical except for the objective-specific instruction.

Specifically, for each user u, the shared context contains:
(1) the user's recent interaction history H_u,
(2) the recalled candidate set C_u ,
(3) candidate-side features, such as item title, price, and popularity-related signals.

In practice, the shared structured context is serialized into a text-based prompt template before being passed to each LLM agent. Each prompt contains two parts: an objective-specific task instruction and the shared structured context that is identical across agents. The shared structured templates and example agent prompts are illustrated in Table 1. For brevity, we show only the Utility Agent prompt as an example. The Revenue Agent and Fairness Agent share the same structure but use different objective-specific instructions: the former emphasizes price-aware monetization, while the latter emphasizes tail-item exposure under relevance constraints. Since current LLM APIs operate over textual inputs, structured recommendation data are transformed into a standardized textual representation that preserves item identity and attribute information while remaining easy for the model to process consistently.

Table 1. Prompt Examples for Multi-Agent Recommendation

Shared Structured Prompt	Utility Agent Prompt
Here are the user interaction history and candidate items. Score this candidate item for your objective. user_id: {*user_id*} preference_summary: {*preference_summary*} recent_history: {*recent_history_text*} candidate_item_id: {*item_id*} candidate_title: {*title*} candidate_category: {*category*} candidate_price: {*price*} candidate_popularity: {*popularity*} base_relevance_score: {*base_score*}...	You are the **Utility Agent** in a multi-agent recommender system.Your task is to estimate how likely the user is to engage with each candidate item. Use both the user's recent-history summary and base relevance score. *Scoring rules:* *- First identify whether the candidate matches the user's recent themes, platforms, device types, or accessory types.* *- If a candidate clearly matches the recent-history summary, it may score above base relevance score.* *- If a candidate is unrelated to the ...*

Multi-agent objective scoring

E-commerce recommendation is inherently multi-objective, as ranking decisions simultaneously affect user preference satisfaction, platform-side value, and supply-side fairness. To explicitly model these heterogeneous objectives, MORE decomposes reranking into three objective-specific scoring processes, each implemented by an LLM-based agent.Let x_u denote the shared structured context for user u, For each candidate item i, the three agents produce objective-specific scores:

$$s_{util}=UtilityAgent(x_u, i)$$

$$s_{rev}=RevenueAgent(x_u, i)$$

$$s_{fair}=FairnessAgent(x_u, i)$$

The **Utility Agent** models the user-side objective and estimates how well each candidate matches the user's current interests. It focuses on the consistency between candidate items and the user's recent interaction history, capturing both semantic similarity and behavioral patterns. The agent is instructed to prioritize preference relevance and avoid being influenced by item price or popularity. As a result, candidates that are strongly aligned with the user's observed interests receive higher utility scores.

The **Revenue Agent** models the platform-side objective by estimating the monetization value of each candidate item. In our setting, this value is primarily reflected by item-side signals such as price. The agent is instructed to favor candidates with higher monetization potential while maintaining basic relevance to the user context. This prevents the model from over-prioritizing high-value but irrelevant items.

The **Fairness Agent** models the supply-side objective by mitigating excessive concentration on head items and improving exposure opportunities for relevant under-represented candidates. It considers popularity-related signals within the candidate set and assigns relatively higher scores to suitable long-tail items when appropriate, while still respecting relevance constraints.

Overall, the three agents provide complementary views of the same reranking instance. Since all agents operate on the same shared structured context, their outputs are directly comparable and can be aggregated in the final reranking stage.

After obtaining the three objective-specific scores, MORE employs a **coordinator** to combine them into a final reranking score:

$$s(u,i)=w_{util} \cdot s_{util}(u,i)+w_{rev} \cdot s_{rev}(u,i)+w_{fair} \cdot s_{fair}(u,i)$$

where $w_{util}, w_{rev}, w_{fair} \geq 0$ denote the aggregation weights for the three objectives.

Experiments

In this section, we evaluate MORE to answer the following research questions:

RQ1: How does MORE compare with conventional rerankers and LLM single-agent reranking?

RQ2: Do the three agents contribute differently to utility, revenue, and fairness?

Setup

We access DeepSeek-V3.2 (Liu, 2024) and Doubao-1.5-pro-32k (Guo, 2025) through API calls with prompt engineering. Experiments are conducted on the Video Games subset of Amazon Reviews'23 (Hou, 2024). We extract user_id, item_id, and timestamp, together with item-side attributes including asin, title, category, and price. During preprocessing, we remove items with missing titles or unparseable prices, retain only interactions that can be mapped to valid items, and filter out users with fewer than five interactions. The remaining interactions are organized into chronological user sequences. Based on the processed data, we randomly sample 10,000 instances for training, and 1,000 instances for testing. Candidate items are first generated through I2I retrieval and long-tail retrieval, and reranking is then performed on the retrieved candidate set.

We compare MORE with the following baselines: I2I Base Order, which directly follows the original retrieval order; GRU4Rec (Hidasi et al., 2015), a representative sequential reranker; Rule-based Multi-Objective Rerank, which replaces LLM-based scoring with manually designed heuristic scores for utility, revenue, and fairness, followed by the same weighted aggregation function; and Single-Agent Multi-Objective, which uses a single LLM to jointly assess all objectives within one prompt and produce a unified reranking score.

We evaluatte multi-objective performance using metrics HitRatio@K (HR@K), NDCG@K, DiscountedRevenue@K (DR@K), AveragePrice@K (AP@K) and TailRatio@K (TR@K), where $K \in \{5,10,20\}$.

Recomendation Quality

User Utility

We evaluate user utility with HR@K and NDCG@K, where HR@K measures whether the target item appears in the top-K list, and NDCG@K further emphasizes higher-ranked hits.As shown in Table 2, the multi-agent method consistently achieves the best user utility performance, outperforming the best

single-agent baseline by up to 33.1% in HR@20 and 24.0% in NDCG@20, while also exceeding the strongest traditional baseline, GRU4Rec, by up to 25.2% and 17.3% on the same metrics.

Table 2. User Utility Evaluation (HR@K and NDCG@K) Across Different Methods
Note: The best results are shown in **bold**, and the second-best results are underlined.

	HR@5	HR@10	HR@20	NDCG@5	NDCG@10	NDCG@20
I2I	0.058	0.081	0.114	0.043	0.050	0.058
GRU4Rec	0.159	0.275	0.456	0.102	0.139	0.185
Rule	0.066	0.153	0.336	0.045	0.073	0.119
Doubao Single-Agent	0.135	0.240	0.429	0.089	0.123	0.170
Doubao Multi-Agents	0.164	0.303	0.555	0.100	0.144	0.208
DeepSeek Single-Agent	0.156	0.260	0.419	0.102	0.136	0.175
DeepSeek Multi-Agents	**0.178**	**0.307**	**0.571**	**0.110**	**0.151**	**0.217**

Table 3. Platform-Side Value (DR@K and AP@K) Across Different Methods

	DR@5	DR@10	DR@20	AP@5	AP@10	AP@20
I2I	2.986	3.514	3.937	51.998	51.450	51.643
GRU4Rec	6.437	8.607	11.235	55.095	57.560	58.078
Rule	5.155	9.416	14.189	**133.641**	**126.796**	**93.034**
Doubao Single-Agent	4.509	6.243	8.863	55.612	51.495	50.126
Doubao Multi-Agents	8.962	12.138	15.490	96.624	83.987	70.407
DeepSeek Single-Agent	6.426	8.023	10.068	60.199	55.995	52.972
DeepSeek Multi-Agents	**10.947**	**13.844**	**17.333**	99.836	90.772	77.128

Platform-Side Value

We evaluate platform-side value using DR@K and AP@K. DR@K measures expected revenue by combining purchase likelihood and item price, while AP@K measures the average price of items.

As shown in Table 3, the multi-agent method achieves the strongest overall platform-side value performance.DeepSeek Multi-Agent delivers the best platform-side value overall, improving over the best single-agent baseline by 70.3% / 72.6% / 72.2% on DR@5/10/20 and 65.9% / 62.1% / 45.6% on AP@5/10/20. While the Rule-based method attains the highest AP@K, it ignores purchase intent and therefore underperforms on DR@K. This shows that multi-agent ranking better balances price promotion and actual purchase likelihood.

Fairness

We evaluate fairness using TR@K, which measures the proportion of long-tail items in the top-K recommendation list. A higher TR indicates better exposure for less popular items and improved supply-side fairness.
AS shown in Table 4, compared with traditional baselines, multi-agent methods remain competitive or superior. While GRU4Rec attains the highest TR@5, its performance drops significantly as K increases, whereas multi-agent methods consistently maintain high TR@10/20.
Overall, the multi-agent approach is close to the Pareto frontier of the three objectives. It not only

maintains competitive user utility, but also secures strong platform-side value and better fairness through enhanced long-tail exposure. These results suggest that the multi-agent framework can effectively balance multiple competing objectives, instead of improving one dimension by sacrificing the others.

Table 4. Fairness (TR@K) Across Different Methods

	TR@5	TR@10	TR@20
I2I	0.176	0.198	0.214
GRU4Rec	**0.461**	0.356	0.271
Rule	0.171	0.184	0.200
Doubao single-agent	0.253	0.259	0.260
Doubao Multi-agents	0.440	**0.414**	**0.404**
DeepSeek single-agent	0.203	0.212	0.218
DeepSeek Multi-agents	0.395	0.397	0.400

Ablation Study

To verify the role of each specialized agent, we remove one agent at a time from the Deepseek Multi-Agent framework and evaluate the corresponding target metric. As shown in Figure 2, when the Utility Agent is removed, NDCG@10 drops from 0.151 to 0.129 (-14.6%). When the Revenue Agent is removed, DR@10 declines from 13.84 to 8.90 (-35.7%). When the Fairness Agent is removed, TR@20 decreases from 0.397 to 0.185 (-53.4%). This confirms that all three agents are effective, with each one primarily contributing to the optimization of its designated objective.

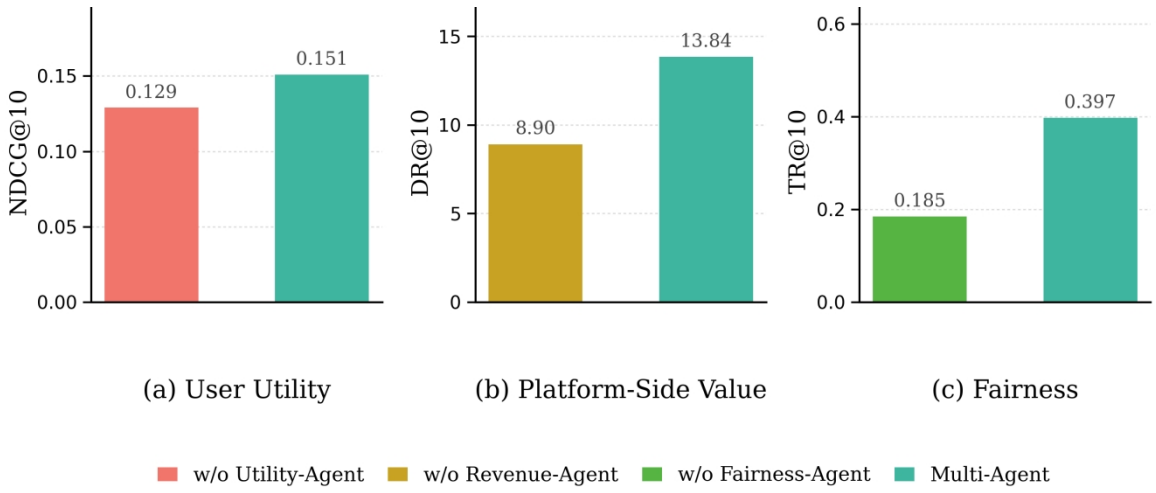


Figure 2. Ablation results by removing each agent from the MORE

Conclusion

In this work, we propose MORE, an LLM-based multi-agent framework for e-commerce recommendation. MORE uses three specialized agents to score candidate items from the user, platform, and supply-side perspectives, and then reranks the candidates by aggregating their scores. Experimental results show that MORE achieves consistent gains on multiple metrics compared with both conventional and single-agent recommenders. We hope this work can offer a useful step toward broader use of LLM agents in e-commerce recommendation.

References

Xiao, B., & Benbasat, I. (2007). E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact1. *MIS quarterly*, 31(1), 137-209.

Lin, J., Dai, X., Xi, Y., Liu, W., Chen, B., Zhang, H., ... & Zhang, W. (2025). How can recommender systems benefit from large language models: A survey. *ACM Transactions on Information Systems, 43*(2), 1-47.

Jia, J., Wang, Y., Li, Y., Chen, H., Bai, X., Liu, Z., ... & Gai, K. (2025, April). LEARN: knowledge adaptation from large language model to recommendation for practical industrial application. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, No. 11, pp. 11861-11869).

Chen, Y., Tan, J., Zhang, A., Yang, Z., Sheng, L., Zhang, E., ... & Chua, T. S. (2024). On softmax direct preference optimization for recommendation. *Advances in Neural Information Processing Systems, 37*, 27463-27489.

Jannach, D., & Abdollahpouri, H. (2023). A survey on multi-objective recommender systems. *Frontiers in big Data, 6,* 1157899.

Jiang, M., Bao, K., Zhang, J., Wang, W., Yang, Z., Feng, F., & He, X. (2024, May). Item-side fairness of large language model-based recommendation system. In *Proceedings of the ACM web conference 2024* (pp. 4717-4726).

Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data mining and knowledge discovery, 5*(1), 115-153.

Wu, X., Chen, H., Zhao, J., He, L., Yin, D., & Chang, Y. (2021, March). Unbiased learning to rank in feeds recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (pp. 490-498).

Sharma, A. K., Bajpai, B., Adhvaryu, R., Pankajkumar, S. D., Gordhanbhai, P. P., & Kumar, A. (2023). An efficient approach of product recommendation system using NLP technique. *Materials Today: Proceedings, 80,* 3730-3743.

Si, Z., Sun, Z., Zhang, X., Xu, J., Zang, X., Song, Y., ... & Wen, J. R. (2023, July). When search meets recommendation: Learning disentangled search representation for recommendation. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval* (pp. 1313-1323).

Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., ... & Li, Q. (2024). Recommender systems in the era of large language models (llms). *IEEE Transactions on Knowledge and Data Engineering, 36*(11), 6889-6907.

Yin, B., Xie, J., Qin, Y., Ding, Z., Feng, Z., Li, X., & Lin, W. (2023, September). Heterogeneous knowledge fusion: A novel approach for personalized recommendation via llm. In *Proceedings of the 17th ACM Conference on Recommender Systems* (pp. 599-601).

Qu, Y., Cai, H., Ren, K., Zhang, W., Yu, Y., Wen, Y., & Wang, J. (2016, December). Product-based neural networks for user response prediction. In *2016 IEEE 16th international conference on data mining (ICDM)* (pp. 1149-1154). IEEE.

Geng, S., Liu, S., Fu, Z., Ge, Y., & Zhang, Y. (2022, September). Recommendation as language processing (rlp): A unified pretrain, personalized prompt & predict paradigm (p5). In *Proceedings of the 16th ACM conference on recommender systems* (pp. 299-315).

Zhang, X., Xin, X., Li, D., Liu, W., Ren, P., Chen, Z., ... & Ren, Z. (2023, February). Variational reasoning over incomplete knowledge graphs for conversational recommendation. In *Proceedings of the sixteenth ACM international conference on web search and data mining* (pp. 231-239).

Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., ... & Piao, Y. (2024). Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.

Guo, D., Wu, F., Zhu, F., Leng, F., Shi, G., Chen, H., ... & Wang, W. (2025). Seed1. 5-vl technical report. *arXiv preprint arXiv:2505.07062*.

Hou, Y., Li, J., He, Z., Yan, A., Chen, X., & McAuley, J. (2024). Bridging language and items for retrieval and recommendation. *arXiv preprint arXiv:2403.03952*.

Hidasi, B., Karatzoglou, A., Baltrunas, L., & Tikk, D. (2015). Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*.

W2

Learning When to Enrich: A Human-AI System for Dynamic Capability Development in Demand Forecasting.

Bongsug Chae (Kansas State University), Eunhye Olivia Park (Gachon University)

Demand forecasting in restaurants often faces resource constraints, including limited analytical talent and technology infrastructure. The academic literature primarily treats demand forecasting as a quantitative problem, employing mathematical and computational approaches that emphasize developing and refining statistical or machine learning (ML) models. While these methods are useful, they rarely view forecasting as a strategic organizational skill that promotes flexibility and resilience. In response, this research offers a broader view by proposing an integrative framework that combines two literatures: dynamic capabilities and foundation models. This study employs design science research (DSR) as its methodology. We are creating an AI system for a large restaurant chain in the United States. Our AI system integrates foundational time-series models with an LLM reasoner into the demand forecasting workflow.

KEY WORDS

Human-AI System, Foundation Model, LLMs, Demand Forecasting, Dynamic Capabilities

KITS·ICEC·KIISS
2026 통합 국제학술대회

W3 Generative AI & LLM Applications

Session Chair : Sung-Byung Yang (Kyung Hee Univ.)



W3

Beyond the Classroom: A Task-Technology Fit Analysis of GenAI Speaking Tutors in Autonomous EFL Learning Completed Research Paper

Pei-Hung Ju, Tzu-Yu Lin, Zhi-Xuan Wu, Shang-Wen Huang, Ting-Chien Chiu
(National Kaohsiung University of Science and Technology),
Hsiao-Lan Wei (National Taiwan University of Science and Technology)

This study applies the Task–Technology Fit (TTF) framework to evaluate how autonomous learners use Generative AI (GenAI) speaking tutors for English as a Foreign Language (EFL) practice outside the traditional classroom. Testing mainstream industry claims against pedagogical realities, we analyzed 308 autonomous EFL learners using PLS–SEM. Results indicate that while Multimodality, Natural Language Processing, Personalized Learning, and Self–Regulated Learning significantly predict Task–Technology Fit and enhance Speaking Performance, Real–Time Feedback and Immersive Learning yield non–significant effects. Exposing a critical mismatch between vendor marketing and autonomous learner realities, these findings highlight the risks of cognitive overload and fragmented, episodic usage. Ultimately, this study recommends optimizing GenAI for short, convenience–driven practice rather than intrusive corrections and prolonged immersion.

KEY WORDS

generative AI, task–technology fit, EFL speaking, self–regulated learning, real–time feedback

W3

Dynamic Modality Weighting with Confidence-Based Fusion in Resource-Constrained Settings

Akshath Hp, Manasa K, Chaitra Reddy B (NItte meenakshi Institue of technlogy)

Multimodal sentiment analysis on resource-constrained devices remains challenging due to the high computational cost of transformer-based approaches. Existing state-of-the-art models such as MulT and MAG-BERT achieve strong accuracy but require GPU inference, millions of parameters, and large labeled datasets, limiting their real- world deployment on edge devices. To address this gap, we propose Dynamic Modality Weighting (DMW), a lightweight○ fusion mechanism using BiGRU encoders that dynamically weights each modality based on learned importance scores. The proposed framework independently estimates scalar importance scores for text, audio, and visual modalities and regulates their contribution through normalized weighting, enabling the system to emphasize informative modalities and suppress noisy inputs without computationally expensive attention layers. Experiments on the CMU-MOSI dataset show that DMW achieves 0.990±0.034 MAE, 74.7±0.8% binary accuracy (negative/non-negative), and 0.746±0.009 F1-score averaged over five random seeds, using only 490K parameters-10× fewer than MulT and 224× fewer than BERT-base. While DMW performs comparably to static fusion baselines on perfectly clean data, single-checkpoint robustness experiments show 6.5-15.2% lower MAE under audio noise and 7.3-7.8% lower MAE under modality dropout compared to the Concatenation baseline, and the framework provides explicit interpretability through learned modality weights. These results offer practitioners a compact, interpretable alternative for edge deployment where computational resources are limited and modality quality varies.

KEY WORDS

Multimodal Sentiment Analysis Dynamic, Modality Weighting Confidence-Based Fusion LowResource, Learning Edge Deployment Robust Multimodal Systems Interpretability

Dynamic Modality Weighting with Confidence-Based Fusion in Resource-Constrained Settings

Akshath H P
akshathhp123@gmail.com
Department of Artificial Intelligence and Machine Learning,
Nitte Meenakshi Institute of Technology,
Nitte (Deemed to be University)
Bengaluru, India

Chaitra Reddy B
chaitrareddy984@gmail.com
Department of Artificial Intelligence and Machine Learning,
Nitte Meenakshi Institute of Technology,
Nitte (Deemed to be University)
Bengaluru, India

Manasa K
mansa.k3030@gmail.com
Department of Artificial Intelligence and Machine Learning,
Nitte Meenakshi Institute of Technology,
Nitte (Deemed to be University)
Bengaluru, India

Sunil Kumar V
sunil.kumar@nmit.ac.in
Department of Artificial Intelligence and Machine Learning,
Nitte Meenakshi Institute of Technology,
Nitte (Deemed to be University)
Bengaluru, India

***Abstract*—Multimodal sentiment analysis on resource-constrained devices remains challenging due to the high computational cost of transformer-based approaches. Existing state-of-the-art models such as MulT and MAG-BERT achieve strong accuracy but require GPU inference, millions of parameters, and large labeled datasets, limiting their real-world deployment on edge devices. To address this gap, we propose Dynamic Modality Weighting (DMW), a lightweight fusion mechanism using BiGRU encoders that dynamically weights each modality based on learned importance scores. The proposed framework independently estimates scalar importance scores for text, audio, and visual modalities and regulates their contribution through normalized weighting, enabling the system to emphasize informative modalities and suppress noisy inputs without computationally expensive attention layers. Experiments on the CMU-MOSI dataset show that DMW achieves 0.990±0.034 MAE, 74.7±0.8% binary accuracy (negative/non-negative), and 0.746±0.009 F1-score averaged over five random seeds, using only 490K parameters—10× fewer than MulT and 224× fewer than BERT-base. While DMW performs comparably to static fusion baselines on perfectly clean data, single-checkpoint robustness experiments show 6.5–15.2% lower MAE under audio noise and 7.3–7.8% lower MAE under modality dropout compared to the Concatenation baseline, and the framework provides explicit interpretability through learned modality weights. These results offer practitioners a compact, interpretable alternative for edge deployment where computational resources are limited and modality quality varies.**

***Index Terms*—Multimodal Sentiment Analysis, Dynamic Modality Weighting, Confidence-Based Fusion, Low-Resource Learning, Edge Deployment, Robust Multimodal Systems, Interpretability**

I. Introduction

Multimodal Sentiment Analysis (MSA) focuses on understanding human sentiment by jointly leveraging information from multiple modalities, commonly text, audio, and visual signals. By integrating linguistic content with acoustic cues and facial expressions, multimodal approaches aim to provide a richer and more accurate interpretation of human emotions compared to unimodal systems. As a result, MSA has gained significant attention in applications such as opinion mining, human–computer interaction, and affective computing.

Despite recent progress, effective multimodal fusion remains a challenging problem. State-of-the-art MSA systems rely on complex architectures such as attention-based fusion mechanisms [1], [2], tensor fusion networks [3], and multimodal transformers to capture cross-modal interactions. While these models achieve strong performance on benchmark datasets, they suffer from high computational complexity (∼5M–110M parameters) and require substantial amounts of labeled data, making them impractical for edge deployment scenarios such as mobile devices, IoT systems, and real-time monitoring applications.

In real-world deployment settings, multimodal data is frequently imperfect. Audio signals may be affected by background noise, recording artifacts, or speaker variability, while visual features may be unreliable due to occlusions, lighting variations, or missing frames. Furthermore, obtaining large-scale labeled multimodal datasets is expensive and time-consuming, particularly for domain-specific applications, low-

resource languages, or privacy-sensitive scenarios such as mental health monitoring. Recent dynamic fusion approaches such as DynMM [5] and Dynamic Attention Fusion [13] (preprint, not peer-reviewed) have addressed adaptive modality selection, but they still rely on complex gating networks or attention mechanisms with significant computational overhead. Confidence-aware frameworks such as FEMF [11] introduce hierarchical prediction with confidence scoring, but the resulting multi-stage architecture increases model complexity beyond what is practical for edge deployment.

This work addresses a fundamentally different problem from accuracy-focused SOTA methods: *How can we deploy multimodal sentiment analysis in resource-constrained, real-world environments with interpretability?* Specifically, we target three critical challenges: (1) *robustness under modality degradation*, (2) *data efficiency in low-resource settings*, and (3) *interpretability of fusion decisions*—while maintaining a computational footprint suitable for CPU-only edge devices.

To this end, we propose a **Dynamic Modality Weighting (DMW)** framework that assigns adaptive importance weights to text, audio, and visual modalities based on their estimated reliability. The technical novelty of our approach lies in a specific two-stage confidence pipeline: a per-modality *sigmoid* head maps each unimodal embedding to a bounded scalar importance score $c_m \in (0, 1)$, which is then jointly normalized via *softmax* across modalities to obtain a valid probability simplex of fusion weights. This formulation differs fundamentally from prior dynamic fusion approaches: DynMM [5] uses a discrete gating network that selects or drops entire modalities and was designed for compute-budgeted inference paths, not interpretability. Zhu et al. [6] estimate uncertainty via Gaussian-assumption variance and applies it only to emotion recognition with a multi-channel LSTM. DAF [13] fuses only two modalities (text + audio) using cross-attention—not a lightweight scalar weighting. None of these provide a *single scalar weight per modality that is simultaneously interpretable, CPU-deployable, and jointly normalized end-to-end without cross-attention*, which is precisely what our sigmoid+softmax formulation achieves.

The proposed framework makes four key contributions:

1) A lightweight architecture employing BiGRU encoders with learned dynamic weighting, requiring only 490K parameters (10× fewer than MulT) while achieving competitive clean-data performance (∼0.990 MAE, 74.7% binary accuracy) against classical static fusion baselines.
2) Systematic evaluation of data efficiency across all lightweight fusion strategies, demonstrating that BiGRU-based architectures achieve practically relevant performance (MAE < 1.3, below the range of the sentiment scale midpoint) with as few as 76 training samples (5% of data).
3) Comprehensive single-checkpoint robustness evaluation demonstrating 6.5–15.2% lower MAE under audio noise and 7.3–7.8% under modality dropout compared to the Concatenation baseline, suggesting the effectiveness of learned importance weighting for unreliable real-world modalities.
4) Explicit interpretability through learned modality weights, providing transparency into fusion decisions—a critical requirement for deployment in sensitive applications such as mental health monitoring and human-computer interaction.

II. Related Work

A. Multimodal Fusion Techniques

Early fusion approaches concatenate features from different modalities before processing. While simple and efficient, these methods treat all modalities equally and fail to account for varying reliability across inputs.

Zadeh et al. [3] introduced the Tensor Fusion Network (TFN), which explicitly models unimodal, bimodal, and trimodal interactions through tensor-based representations. While TFN achieved breakthrough performance, it suffers from exponential memory growth ($O(d^3)$), making it impractical for resource-constrained deployments. Liu et al. [7] proposed Low-rank Multimodal Fusion (LMF) to address TFN's scalability issues through low-rank tensor decomposition, representing an early effort toward efficient multimodal fusion.

Tsai et al. [1] proposed the Multimodal Transformer (MulT), applying cross-modal attention mechanisms to align unaligned sequences across modalities. MulT demonstrated strong performance on CMU-MOSI (MAE ≈ 0.861) and CMU-MOSEI benchmarks. However, the multi-head attention layers introduce significant computational overhead (∼5M parameters, ∼2.5G FLOPs), requiring GPU acceleration and substantial training data.

Rahman et al. [8] introduced MAG-BERT, integrating multimodal adaptation gates with pre-trained BERT representations. While achieving state-of-the-art accuracy, MAG-BERT requires 110M+ parameters and cloud-scale infrastructure, limiting practical deployment.

Cai et al. [17] proposed a multi-layer feature fusion framework combined with multi-task learning for MSA, demonstrating that hierarchical feature representations across modalities improve sentiment prediction. While their approach achieves strong performance, the multi-layer fusion architecture and multi-task training pipeline introduce additional computational overhead compared to our single-stage confidence-driven fusion strategy.

B. Dynamic and Confidence-Based Fusion

Xue and Marculescu [5] proposed DynMM, a dynamic fusion framework with gating networks for adaptive modality selection. While conceptually similar to our approach in adapting per-sample, DynMM employs more complex gating mechanisms and does not focus on the efficiency-robustness trade-off or data efficiency analysis that forms our primary contribution.

Zhu et al. [6] proposed a dynamic confidence-aware multimodal framework for emotion recognition that adaptively

weighs each modality based on its reliability. Their work validates the confidence-based paradigm but targets a different task domain with different architectural choices.

The FEMF framework [11] introduces confidence-aware fusion with hierarchical prediction and frequency-aware experts. While effective, their multi-stage architecture is significantly more complex than our single-stage confidence estimation, demonstrating the need for simpler alternatives in resource-constrained settings.

Hazarika et al. [2] presented MISA, which decomposes representations into modality-invariant and modality-specific components. While effective, MISA's dual-encoder architecture increases complexity and training time, limiting applicability in low-resource scenarios.

Nagrani et al. [9] introduced attention bottlenecks to control information flow during multimodal fusion, demonstrating that selective interaction can improve generalization—a principle that supports our lightweight weighting approach.

C. Low-Resource and Efficient Multimodal Learning

Recent work has begun addressing multimodal learning with limited data. Prompt-based methods leverage large language models for few-shot MSA, but require access to billion-parameter foundation models. Yu et al. [10] proposed Self-MM for self-supervised multi-task learning in MSA, showing data-efficient strategies through auxiliary tasks. Similarly, Cai et al. [17] incorporate multi-task learning objectives to improve generalization under limited supervision; however, both Self-MM and this approach still require substantial computational resources for their multi-task training pipelines, limiting their applicability in resource-constrained edge settings.

GRU-based lightweight architectures remain relevant for efficient multimodal processing, as demonstrated by Qin et al. [12], who propose MulG—a model combining GRU encoders with directed pairwise cross-modal attention. MulG achieves strong performance (82.2% Acc-2 on CMU-MOSI) by leveraging cross-modal attention to synchronize asynchronous signals. However, the cross-modal attention layers add significant parameters compared to our attention-free approach, and MulG does not study robustness under modality degradation or data efficiency—gaps our work addresses.

Comprehensive surveys by Zhu et al. [14] and Dao et al. [15] highlight that lightweight and interpretable fusion strategies remain underexplored, particularly in low-resource and edge deployment settings—the exact gap our work addresses.

D. Our Contribution

Unlike prior work that prioritizes raw performance through complex architectures, our DMW framework uniquely combines *computational efficiency*, *robustness analysis*, *data efficiency evaluation*, and *explicit interpretability* in a single lightweight system. While DynMM [5] and DAF [13] (preprint) explore dynamic fusion, they do not systematically study the efficiency-robustness-interpretability trade-off that is central to our contribution. Our specific formulation using sigmoid confidence heads with softmax normalization provides per-sample interpretable weights without cross-modal attention overhead, achieving competitive performance with $10\times$ fewer parameters than transformer-based alternatives.

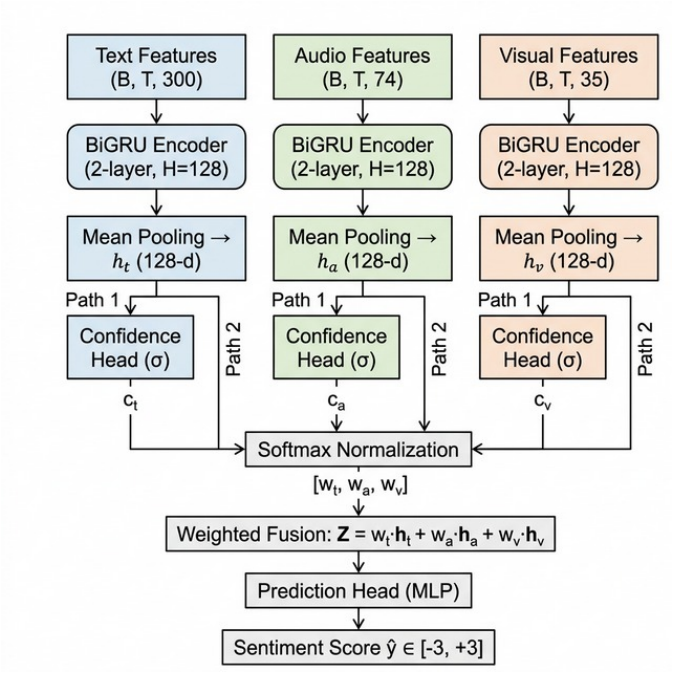


Fig. 1. Architecture of the proposed DMW framework. Parallel BiGRU encoders extract modality-specific embeddings. Per-modality importance heads produce scalar scores normalized via softmax, yielding dynamic weights for weighted fusion.

III. Proposed Method

This section presents the proposed Dynamic Modality Weighting (DMW) framework for multimodal sentiment analysis. The method addresses challenges arising from varying modality reliability, limited training data, and practical deployment constraints through a lightweight, confidence-driven fusion strategy.

A. Method Overview

The proposed DMW framework processes text, audio, and visual inputs through parallel unimodal encoding pipelines to extract modality-specific representations. A confidence score is estimated for each modality to assess its reliability for the given input. These confidence scores are normalized to obtain dynamic modality weights, which fuse the unimodal representations through weighted summation. The resulting fused representation is passed to a regression head to predict a continuous sentiment score. Fig. 1 illustrates the complete architecture.

B. Dataset Description

The proposed framework is evaluated on the CMU-MOSI (Multimodal Opinion Sentiment Intensity) dataset [4], which consists of 2,199 opinion-level video segments from 93 YouTube speakers annotated with continuous sentiment intensity scores in the range $[-3, +3]$. Each sample contains

temporally aligned text, audio, and visual modalities. We use the standard split: 1,538 training, 220 validation, and 686 test samples.

Pre-extracted features provided with the CMU-MOSI dataset are used for all modalities:

- **Text**: 300-dimensional GloVe word embeddings [16], consistent with prior lightweight baselines (TFN [3], LMF [7]), enabling fair comparison without introducing large language model dependencies.
- **Audio**: 74-dimensional COVAREP audio descriptors encoding prosodic and spectral characteristics.
- **Visual**: 35-dimensional FACET features capturing facial expression-related action units.

To evaluate data efficiency, we conduct experiments with varying training data fractions $f \in \{0.05, 0.10, 0.20, 0.50, 1.0\}$, simulating real-world scenarios where labeled data is scarce.

The CMU-MOSI dataset is publicly available at http://multicomp.cs.cmu.edu/resources/cmu-mosi-dataset/.

C. Unimodal Feature Extraction

Each modality is processed independently using a Bidirectional Gated Recurrent Unit (BiGRU) network to capture temporal dependencies within the input sequences. Global Mean Pooling converts variable-length hidden state sequences into fixed-length embeddings:

$$\mathbf{h}_m = \text{MeanPool}(\text{BiGRU}(\mathbf{X}_m)), \quad m \in \{t, a, v\} \tag{1}$$

where $\mathbf{h}_m \in \mathbb{R}^H$ with hidden dimension $H = 128$. Each BiGRU uses 2 layers with 64 hidden units per direction and dropout rate 0.3.

D. Dynamic Modality Importance Weighting

An importance estimation head is applied to each modality-specific embedding to learn its relative contribution to the final prediction:

$$c_m = \sigma(\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \mathbf{h}_m + b_1) + b_2) \tag{2}$$

where $\mathbf{W}_1 \in \mathbb{R}^{64\times128}$, $\mathbf{W}_2 \in \mathbb{R}^{1\times64}$ are learnable parameters, and $\sigma(\cdot)$ denotes the sigmoid function. The importance score $c_m \in (0, 1)$ is learned end-to-end through the sentiment prediction objective. While trained without explicit noise supervision, the learned weights implicitly reflect statistical importance patterns in the training data, enabling graceful degradation when modality quality varies at inference time.

The importance scores are normalized using softmax to obtain dynamic modality weights:

$$[w_t, w_a, w_v] = \text{Softmax}([c_t, c_a, c_v]) \tag{3}$$

These weights satisfy $w_t + w_a + w_v = 1$ and dynamically control each modality's influence during fusion.

TABLE I
FUNCTIONAL LAYERS OF THE PROPOSED FRAMEWORK

Component	Layer	Functional Description
Text Encoder	Encoding	BiGRU (2-layer, 64×2) + mean pooling
Audio Encoder	Encoding	BiGRU (2-layer, 64×2) + mean pooling
Visual Encoder	Encoding	BiGRU (2-layer, 64×2) + mean pooling
Importance Module	Control	Linear(128→64→1) + Sigmoid
Weighting Module	Control	Softmax normalization for adaptive weights
Fusion Module	Control	Weighted summation
Prediction Head	Output	Linear(128→64→1) for regression

E. Weighted Multimodal Fusion

The unimodal embeddings are combined through weighted summation:

$$\mathbf{Z} = w_t \mathbf{h}_t + w_a \mathbf{h}_a + w_v \mathbf{h}_v \tag{4}$$

This preserves the original embedding dimensionality ($\mathbb{R}^{128}$) while maintaining explicit interpretability of each modality's contribution.

F. Sentiment Prediction

The fused representation is passed to a prediction head:

$$\hat{y} = \mathbf{W}_p(\text{ReLU}(\mathbf{W}_f \mathbf{Z} + b_f)) + b_p \tag{5}$$

The model is trained end-to-end using Mean Squared Error (MSE) loss:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{6}$$

The model is optimized with Adam (learning rate 10^{-3}, batch size 32, 20 epochs) and gradient clipping (max norm 1.0). Dropout (0.3) is applied for regularization.

G. Functional Components

Table I summarizes the key functional components.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

All experiments are conducted on the CMU-MOSI dataset using the standard train/validation/test split. All models are implemented in PyTorch 2.1 and trained using the Adam optimizer with learning rate 10^{-3}, batch size 32, and 20 training epochs. The model checkpoint achieving the lowest validation MAE is selected for test set evaluation. The five seeds used for multi-seed experiments are $\{42, 123, 456, 789, 1024\}$. Latency measurements were performed on an Intel Core i5-1135G7 CPU @ 2.40 GHz with 8 GB RAM running 64-bit Windows. Code will be released upon acceptance. Model performance is evaluated using four standard metrics:

TABLE II
COMPLETE RESULTS ON CMU-MOSI TEST SET (MEAN ± STD OVER 5 SEEDS). ACC-2 USES NEGATIVE/NON-NEGATIVE CONVENTION (THRESHOLD AT 0), EXCLUDING EXACT ZEROS.

Model	MAE ↓	Acc-2 ↑	F1 ↑	Corr ↑
DMW (Ours)	0.990±0.034	0.747±0.008	0.746±0.009	0.626±0.016
Concatenation	**0.964±0.008**	**0.757±0.007**	**0.756±0.007**	**0.640±0.009**
Fixed-Weight	0.970±0.025	0.755±0.010	0.754±0.010	0.632±0.012
TFN [3]	0.970	0.736	–	0.633
LMF [7]	0.912	0.760	–	0.668
MulG [12]	–	0.822	–	–
MulT [1]	0.861	0.837	0.836	0.711
MAG-BERT [8]	∼0.73	0.842	–	0.796

- **MAE** (Mean Absolute Error, ↓): Primary regression metric
- **Acc-2** (Binary Accuracy, ↑): Negative/non-negative classification (threshold at 0, excluding exact zeros)
- **F1** (Weighted F1-score, ↑): Harmonic mean of precision and recall
- **Corr** (Pearson Correlation, ↑): Linear relationship quality

No formal statistical significance tests (e.g., paired t-test) were applied; reported differences under 3% between lightweight models should be interpreted cautiously as they may fall within the range of seed-to-seed variability.

B. Baseline Models

DMW is compared against two baseline fusion approaches and state-of-the-art reference models:

1) **Concatenation Baseline**: Unimodal embeddings are concatenated ($\mathbb{R}^{3H}$) and passed to a regression head.
2) **Fixed-Weight Baseline**: Unimodal embeddings are combined using uniform weights ($w = 1/3$) without confidence estimation.

All baselines utilize identical unimodal encoders, ensuring observed differences are attributable solely to the fusion strategy. We note that a unimodal text-only baseline was not included; given text's dominant average weight (0.42, see Section IV-H), future work should verify that multimodal fusion provides gains over strong unimodal text models.

C. Performance with Full Metrics

Table II presents the complete evaluation results using all standard MOSI metrics, addressing the requirement for comprehensive evaluation beyond MAE alone.

Key observations from Table II:

1) **Comparable performance on clean data**: Averaged over 5 runs, DMW achieves performance numerically close to both Concatenation and Fixed-Weight baselines (0.990 MAE vs 0.964/0.970). While Concatenation holds a slight edge on perfectly clean benchmark data, DMW's true advantage lies in robustness to noise (as detailed in Section IV-G).
2) **Competitive with baseline TFN**: DMW's clean-data MAE (0.990) and Acc-2 (74.7%) are directly comparable to the classical Tensor Fusion Network (0.970 MAE, 73.6% Acc-2), while utilizing 4× fewer parameters, demonstrating strong baseline efficiency.
3) **Performance gap vs attention-based models reflects architectural trade-off**: MulT (0.861 MAE) and MulG (82.2% Acc-2) achieve better performance through cross-modal attention, but require 10× more parameters (MulT) or cross-modal attention overhead (MulG). Our contribution targets a complementary problem—lightweight, interpretable, CPU-deployable sentiment analysis without cross-modal attention.

TABLE III
DATA EFFICIENCY: PERFORMANCE VS TRAINING DATA SIZE. ROWS FOR 5%–50% DATA FRACTIONS USE A SINGLE TRAINING RUN. THE 100% ROW USES 5-SEED AVERAGES (MEAN±STD, IDENTICAL TO TABLE II). NO STD DEVS AVAILABLE FOR SINGLE-RUN ROWS.

Data	N	Model	MAE↓	Acc2↑	F1↑
5%	76	DMW	**1.280**	0.649	0.644
		Concat	1.298	0.647	0.638
		Fixed	1.297	**0.656**	**0.656**
10%	153	DMW	1.156	0.687	0.686
		Concat	1.158	0.680	0.680
		Fixed	**1.132**	**0.711**	**0.710**
20%	307	DMW	1.090	0.708	0.704
		Concat	1.056	0.722	0.722
		Fixed	**1.046**	**0.739**	**0.738**
50%	769	DMW	1.029	0.734	0.733
		Concat	1.027	**0.748**	**0.745**
		Fixed	**1.024**	0.734	0.734
100%	1538	DMW	0.990±0.034	0.747±0.008	0.746±0.009
		Concat	**0.964±0.008**	**0.757±0.007**	**0.756±0.007**
		Fixed	0.970±0.025	0.755±0.010	0.754±0.010

D. Data Efficiency Analysis

Table III presents performance across training data fractions with all standard metrics.

Key findings:

1) **All lightweight architectures are data-efficient**: All three fusion strategies achieve practically relevant performance (MAE < 1.3, below the range of the sentiment scale midpoint) with just 76 training samples (5%), suggesting that BiGRU-based architectures are inherently data-efficient regardless of fusion mechanism. However, as these low-data results are based on single runs, the observed differences should be treated as indicative rather than definitive.
2) **Static fusion appears to excel in low data regimes**: At 10–50% data fractions, Fixed-Weight and Concatenation generally achieve better MAE than DMW in these single-run experiments. This pattern suggests that their architectural simplicity may avoid overfitting when limited data prevents the dynamic weighting mechanism from learning meaningful importance patterns.
3) **Dynamic weighting benefits from data scaling**: DMW shows steady improvement with scale, narrowing but not fully closing the gap with static baselines on clean

TABLE IV
EFFICIENCY-PERFORMANCE TRADE-OFF ANALYSIS

Model	Params	FLOPs[a]	Latency	MAE
DMW (Ours)	490K	60.9M	46ms	0.990
Concat	481K	60.9M	50ms	0.964
Fixed-Weight	465K	60.9M	48ms	0.970
TFN [3]	2.0M	500M	–	0.970
LMF [7]	1.2M	300M	–	0.912
MulT [1]	5.0M	2.5G	GPU	0.861
MAG-BERT [8]	110M	22G	Cloud	∼0.73

[a] FLOPs are identical across DMW, Concat, and Fixed-Weight because over 99% of computation resides in the shared BiGRU encoders; the fusion layers (importance heads, softmax, concatenation) contribute negligible FLOPs (<0.1M).

data (0.990 vs 0.964 MAE at 100%). This suggests that learning instance-specific modality importance requires sufficient labeled data to converge effectively, and the dynamic weighting overhead does not translate to clean-data accuracy gains.

E. Efficiency-Performance Trade-off

Table IV presents the Pareto efficiency analysis, showing the trade-off between model complexity and performance. This is a critical analysis for deployment-focused applications.

Key findings:

1) **Parameter Efficiency**: DMW requires only 490K parameters—10× reduction vs MulT (5M) and 224× vs BERT-based models (110M).
2) **FLOPs Reduction**: 60.9M FLOPs represents 41× reduction vs MulT (2.5G) and 361× vs BERT (22G).
3) **Real-Time CPU Inference**: 46ms average latency (measured as single-sample inference, batch size = 1) enables real-time processing at ∼21 samples/second on CPU without GPU.
4) **Comparable to TFN**: DMW achieves comparable MAE (0.990 vs 0.970) to TFN with 4× fewer parameters, demonstrating that importance-based fusion provides a competitive efficiency-accuracy trade-off relative to tensor-based fusion.

F. Fusion Strategy Comparison

To compare fusion strategies under controlled conditions, we evaluate three variants that share identical BiGRU unimodal encoders and differ only in how modality representations are combined.

The comparison, averaged over 5 random seeds to ensure statistical significance, indicates that learned importance weighting does not explicitly improve raw prediction accuracy on perfectly clean benchmark test data compared to simple feature concatenation. In fact, concatenation holds a marginal ∼1–3% edge across metrics. This suggests that when all modalities are of high quality, a linear projection of concatenated features is highly effective. However, this clean-data performance sets the stage for our central contribution: as demonstrated in the following section, static fusion models degrade more rapidly when modality quality varies, whereas DMW's learned weights may help preserve predictive integrity.

TABLE V
FUSION STRATEGY COMPARISON: MEAN METRICS ACROSS 5 RUNS. ALL MODELS SHARE IDENTICAL BIGRU ENCODERS.

Configuration	MAE↓	Acc2↑	F1↑	Corr↑
DMW (Learned Wt.)	0.990	0.747	0.746	0.626
Fixed Equal Wt.	0.970	0.755	0.754	0.632
Concatenation (no wt.)	**0.964**	**0.757**	**0.756**	**0.640**
Δ *DMW vs Concat (negative = DMW worse):*				
	−2.7%	−1.3%	−1.3%	−2.2%

TABLE VI
ROBUSTNESS TO AUDIO NOISE (MAE ↓). RESULTS FROM A SINGLE CHECKPOINT PER MODEL (SEED 42); TRENDS ARE INDICATIVE, NOT STATISTICALLY CONCLUSIVE.

Noise	DMW	Concat	Fixed	DMW Gain[a]
Clean	0.960	1.005	0.963	–
$\sigma = 0.1$	1.01	1.08	1.02	+6.5%
$\sigma = 0.3$	1.06	1.19	1.13	+10.9%
$\sigma = 0.5$	1.12	1.32	1.24	+15.2%

[a] Gain $= (\text{MAE}_{\text{Concat}} - \text{MAE}_{\text{DMW}})/\text{MAE}_{\text{Concat}} \times 100$

G. Robustness Analysis

To evaluate robustness under modality degradation, we conducted two systematic experiments using a single trained checkpoint per model (random seed 42). As these results are based on a single checkpoint rather than multi-seed averages, they should be interpreted as indicative of robustness trends rather than statistically conclusive. The clean-data baselines for this checkpoint—DMW: 0.960 MAE, Concat: 1.005, Fixed: 0.963—differ slightly from the 5-seed averages in Table II (DMW: 0.990, Concat: 0.964, Fixed: 0.970), which is expected due to seed variance. All degradation measurements in Tables VI and VII are relative to the *same* single checkpoint, ensuring that DMW Gain values reflect fair within-checkpoint comparisons and are not confounded by cross-seed variability.

Audio Noise Injection: We added Gaussian noise with varying standard deviations ($\sigma \in \{0.1, 0.3, 0.5\}$) to audio features during inference.

Modality Dropout: We simulated missing modalities by zeroing feature vectors during inference.

These single-checkpoint results suggest that DMW maintains lower MAE than the Concatenation baseline under degradation: 6.5–15.2% lower under audio noise (Table VI) and 7.3–7.8% lower under modality dropout (Table VII). We hypothesize that the learned importance weights implicitly reduce reliance on degraded modalities, though direct weight logging under noise conditions is left for future verification. This capability, if confirmed at scale, would be valuable for real-world deployments where sensor quality varies unpredictably.

3) **Cross-lingual generalization**: Evaluation on low-resource languages and cross-lingual transfer scenarios remains unexplored.
4) **Hardware validation**: Deployment latency should be validated on actual edge hardware such as mobile SoCs and Raspberry Pi.
5) **Self-supervised importance estimation**: Pre-training the importance heads with self-supervised objectives could reduce dependence on labeled data.
6) **Limited robustness evaluation**: Robustness experiments (Tables VI–VII) use a single trained checkpoint. Future work should average robustness metrics over multiple seeds for statistical reliability.

REFERENCES

[1] Y.-H. H. Tsai, S. Bai, P. P. Liang, J. Z. Kolter, L.-P. Morency, and R. Salakhutdinov, "Multimodal transformer for unaligned multimodal language sequences," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2019, pp. 6558–6569.

[2] D. Hazarika, R. Zimmermann, and S. Poria, "MISA: Modality-invariant and -specific representations for multimodal sentiment analysis," in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 1122–1131.

[3] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency, "Tensor fusion network for multimodal sentiment analysis," in *Proc. Conf. Empirical Methods Natural Language Process. (EMNLP)*, 2017, pp. 1103–1114.

[4] A. Zadeh, R. Zellers, E. Pincus, and L.-P. Morency, "Multimodal sentiment intensity analysis in videos: Facial gestures and verbal messages," *IEEE Intell. Syst.*, vol. 31, no. 6, pp. 82–88, 2016.

[5] Z. Xue and R. Marculescu, "Dynamic multimodal fusion," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognition Workshops (CVPRW)*, 2023.

[6] Q. Zhu, C. Zheng, Z. Zhang, W. Shao, and D. Zhang, "Dynamic confidence-aware multi-modal emotion recognition," *IEEE Trans. Affective Comput.*, vol. 15, no. 3, pp. 1358–1370, 2023.

[7] Z. Liu, Y. Shen, V. B. Lakshminarasimhan, P. P. Liang, A. Zadeh, and L.-P. Morency, "Efficient low-rank multimodal fusion with modality-specific factors," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2018, pp. 2247–2256.

[8] W. Rahman, M. K. Hasan, S. Lee, A. Zadeh, C. Mao, L.-P. Morency, and E. Hoque, "Integrating multimodal information in large pretrained transformers," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2020, pp. 2359–2369.

[9] A. Nagrani, S. Yang, A. Arnab, A. Jansen, C. Schmid, and C. Sun, "Attention bottlenecks for multimodal fusion," in *Advances in Neural Information Processing Systems*, vol. 34, 2021, pp. 14200–14213.

[10] H. Yu, H. Xu, P. Kou, J. Du, and L. Zhou, "Learning modality-specific representations with self-supervised multi-task learning for multimodal sentiment analysis," in *Proc. 35th AAAI Conf. Artif. Intell.*, vol. 35, no. 12, 2021, pp. 10790–10797.

[11] X. Zhu and Y. Li, "Frequency-aware experts with multi-stage fusion for multimodal sentiment analysis," *J. Intell. Inf. Syst.*, 2025 (in press).

[12] Z. Qin, Q. Luo, Z. Zang, and H. Fu, "Multimodal GRU with directed pairwise cross-modal attention for sentiment analysis," *Sci. Rep.*, vol. 15, art. no. 10023, Mar. 2025.

[13] S. Abdulhalim, M. Albaghdadi, and M. Farazi, "Multi-modal sentiment analysis with dynamic attention fusion," *arXiv preprint arXiv:2509.22729*, 2025 (preprint, not peer-reviewed).

[14] L. Zhu, Z. Zhu, C. Zhang, Y. Xu, and X. Kong, "Multimodal sentiment analysis based on fusion methods: A survey," *Inf. Fusion*, vol. 95, pp. 306–325, 2023.

[15] P. Q. Dao et al., "Exploring multimodal sentiment analysis models: A comprehensive survey," in *Proc. IEEE Int. Conf. Multimedia Analysis Pattern Recogn. (MAPR)*, 2024.

[16] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Language Process. (EMNLP)*, 2014, pp. 1532–1543.

[17] Y. Cai, X. Li, Y. Zhang, J. Li, F. Zhu, and L. Rao, "Multimodal sentiment analysis based on multi-layer feature fusion and multi-task learning," *Sci. Rep.*, vol. 15, no. 1, art. no. 2126, 2025.

W3

Examining the Developmental Trajectory of AI in Consumer Research: A Citation Network Analysis.

Ching Yung Lin, Chen-Hao Huang, Tai-Yu Lin
(Department of Business Administration National Cheng Kung University).

This study examines the developmental trajectory of AI in consumer related research. Using 961 articles published between 2016 and 2025 in the WOS SSCI and SCI-Expanded databases, this study applies key-route main path analysis and multiple global main path analysis to trace the evolution of the field. The results identify three phases: Perceptual Appraisal, in which AI is viewed as an object of consumer evaluation; Interactional Interpreting, in which AI functions as an interactive mechanism shaping decision-making and experience; and Adaptive Persuasion, in which AI becomes a persuasive and adaptive force through AI-generated content, disclosure, and mediated presentation. In addition, the multiple global main path analysis reveals the AI-authorship effect as the dominant contemporary research theme. This study contributes by providing a structured evolutionary framework for understanding AI in consumer research.

KEY WORDS

AI, Article Intelligence, Consumer, Main Path Analysis, Systematic Review

W3

From Text to Presence: Mechanisms of Generative AI Comments in Social Media Engagement

Yijun Yan, Xusen Cheng, Changyu Zhou (Renmin University of China)

Generative AI (GAI) is transforming online social interactions. While prior research focuses on information quality and anthropomorphic similarity, the role of AI expression style remains underexplored. This study examines how three textual features of GAI comments (algorithmic absurdity, emotional consistency, and topic relevance) influence interaction intentions through social presence, with users' value needs (utilitarian vs. hedonic) as moderators. We introduce algorithmic absurdity as a novel stylistic construct rooted in the structural properties of large language models, distinct from human humor. Using a mixed-methods design, we develop and validate a measurement scale for algorithmic absurdity and empirically test our model. Our findings offer new insights into how GAI reshapes human-AI social interaction through language style.

KEY WORDS

Generative AI, social media, algorithmic absurdity, social presence

From Text to Presence: Mechanisms of Generative AI Comments in Social Media Engagement

Yan, Yijun[a], Cheng, Xusen[b] and Zhou, Changyu[c]
[a] School of Information, Renmin University of China
E-mail: yanyijun@ruc.edu.cn
[b] School of Information, Renmin University of China
E-mail: xusen.cheng@ruc.edu.cn
[c] School of Information, Renmin University of China
E-mail: zcy0626@ruc.edu.cn

Abstract

Generative AI (GAI) is transforming online social interactions. While prior research focuses on information quality and anthropomorphic similarity, the role of AI expression style remains underexplored. This study examines how three textual features of GAI comments (algorithmic absurdity, emotional consistency, and topic relevance) influence interaction intentions through social presence, with users' value needs (utilitarian vs. hedonic) as moderators. We introduce algorithmic absurdity as a novel stylistic construct rooted in the structural properties of large language models, distinct from human humor. Using a mixed-methods design, we develop and validate a measurement scale for algorithmic absurdity and empirically test our model. Our findings offer new insights into how GAI reshapes human-AI social interaction through language style.

Keywords: Generative AI, social media, algorithmic absurdity, social presence

Introduction

Generative artificial intelligence (GAI) has rapidly transformed online communication (Gao et al., 2025). While prior research predominantly emphasizes AI's information quality and anthropomorphic similarity (Cheng et al., 2022; Qiu and Benbasat, 2009), it often overlooks the intrinsic textual features generated by the LLM itself. In social media interactions, the expression style is as critical as the content in shaping user engagement. Therefore, we focus on how GAI's unique stylistic cues affect users' interaction intentions.

To fill this gap, we introduce a concept of "algorithmic absurdity", which refers to statements that deviate from the expected conventional semantics or slightly violate social norms, thereby creating a sense of incongruity. Algorithmic absurdity stems from the structural characteristics of language models. We consider algorithmic absurdity, along with emotional consistency and topic relevance, as inherent features of GAI comments. We study how these features interact with users' value needs (utilitarian and hedonic) to influence the perceived social presence and drive interaction intentions.

Conceptual Development

Incongruity Theory

Incongruity theory posits that humor arises from the perceived gap between expectations and reality. When GAI generates algorithmically absurd responses, it triggers a cognitive process of incongruity resolution. Users first experience a semantic shock, then resolve it by attributing a playful digital

persona to the AI. This resolution process elevates social presence by transforming technical deviations into meaningful social cues.

AI Comment Characteristics, User Value Needs and Social Presence

Drawing on qualitative interviews, we identify three core dimensions of AI comment characteristics: algorithmic absurdity, emotional consistency, and topic relevance. Algorithmic absurdity refers to AI-generated expressions that are unexpected or deviate from conventional norms, producing a subtle sense of incongruity in conversation. Such deviation reflects playful uniqueness rather than system error. Emotional consistency refers to the extent to which AI comments match the emotional tone of the original post. Topic relevance reflects how closely a comment relates to the ongoing discussion in terms of content and context.

In IS research, users' behavior is often shaped by their intrinsic values, cognition, and motivational needs (Quinn and Gutt, 2025). On social media, users pursue various value-based goals through engagement. Based on interviews, we identified two key value dimensions: utilitarian and hedonic. Utilitarian value reflects users' expectations of practical usefulness in online interactions. Hedonic value signifies the pursuit of pleasure and entertainment in the digital environment.

Social presence is the feeling of being with another. In mediated communication, it refers to the sense of being with another human being enabled through technology and the extent to which a person is perceived as a "real person" (Biocca et al., 2003). Research indicates that high message interactivity can compensate for a lack of visual cues and enhances social presence (Go and Sundar, 2019). We contend that AI comment features and users' value jointly shape social presence.

Hypotheses

H1: The algorithmic absurdity of AI comments has an inverted U-shaped relationship with social presence.

H2: The emotional consistency of AI comments positively affects users' social presence.

H3: The topic relevance of AI comments positively affects users' social presence.

H4: Social presence acts as a mediator between the characteristics of AI comments and interaction intention.

H5a: Hedonic value need moderates the inverted U-shaped relationship between algorithmic absurdity and social presence, such that the effect is more pronounced for users with higher hedonic value need.

H5b: Hedonic value need moderates the positive relationship between emotional consistency and social presence, such that the effect is stronger for users with higher hedonic value need.

H5c: Utilitarian value need moderates the positive relationship between topic relevance and social presence, such that the effect is stronger for users with higher utilitarian value need.

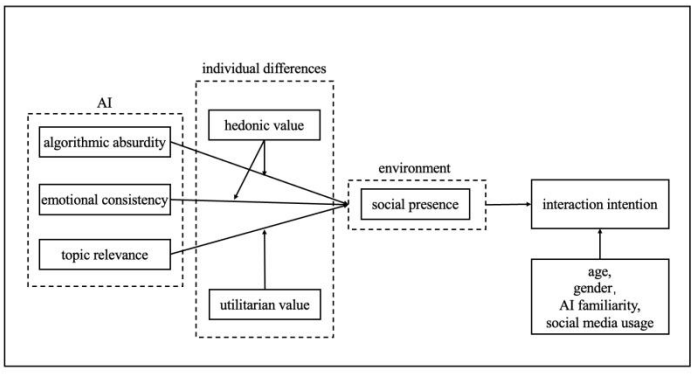


Figure 1. Research Model

Research Methodology

In this study, we used a hybrid approach with two phases: qualitative and quantitative. This design leverages the strengths of both methods to enhance robustness. The qualitative phase enables

exploratory investigation and identification of context-specific constructs. The quantitative phase tests the proposed framework empirically (Venkatesh et al., 2012).

Phase 1: Qualitative Analysis

In the first phase, we recruited 30 social media users who were familiar with GAI comments for semi-structured interviews. Each interview lasted 15 to 20 minutes. These interviews helped identify key constructs: the features of GAI comments (algorithmic absurdity, emotional consistency, and topic relevance), users' value needs (utilitarian and hedonic), and the main mediating variable (social presence). The first phase interview results provided a basis for developing measurement indicators and the structural equation model.

Phase 2: Quantitative Analysis

In the second phase, we designed a structured survey to test the conceptual framework. First, the algorithmic absurdity scale was newly developed based on interview insights and refined through pre-testing. Both Cronbach ' s alpha and composite reliability (CR) of the initial scale exceeded recommended thresholds, confirming its reliability. Second, all latent variables were measured using a 7-point Likert scale. The formal research data were analyzed using structural equation modeling (SEM) for path analysis. Preliminary SEM results supported the hypothesized relationships.

Potential Contributions

Can the comments generated by GAI promote the active participation of platform users? Which features of GAI meet the value demands of users? When GAI begins to have "personality", how do humans perceive and respond to this non-human social style? To address these questions, our contributions are twofold. First, we introduce and formally define algorithmic absurdity and develop a scale with verified reliability. Second, by identifying AI comment style as a driver of social presence in human-AI interaction, we advance social presence theory. These contributions provide a new IS perspective on how GAI reshapes social interactions through language.

Acknowledgements

Xusen Cheng is the corresponding author. We would like to thank the National Natural Science Foundation of China (Grant No. 72434006), Key Projects of Philosophy and Social Sciences Research of Chinese Ministry of Education (Grant No.23JZD022) and National Natural Science Foundation of China (Grant No. 72271236) for providing funding for this research.

References

Biocca, F., Harms, C., and Burgoon, J. K. 2003. "Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria," *Presence: Teleoperators and Virtual Environments* (12:5), MIT Press - Journals, pp. 456–480.

Cheng, X., Zhang, X., Cohen, J., and Mou, J. 2022. "Human vs. AI: Understanding the Impact of Anthropomorphism on Consumer Response to Chatbots from the Perspective of Trust and Relationship Norms," *Information Processing & Management* (59:3), p. 102940.

Gao, Y., Zhang, M. M., and Lysyakov, M. 2025. "Does Social Bot Help Socialize? Evidence from a Microblogging Platform," *Information Systems Research*, INFORMS. https://doi.org/10.1287/isre.2024.1089.

Go, E., and Sundar, S. S. 2019. "Humanizing Chatbots: The Effects of Visual, Identity and Conversational Cues on Humanness Perceptions," *Computers in Human Behavior* (97), Elsevier BV, pp. 304–316.

Qiu, L., and Benbasat, I. 2009. "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems," *Journal of Management Information Systems* (25:4), Informa UK Limited, pp. 145–182.

Quinn, M., and Gutt, D. 2025. "Heterogeneous Effects of Generative Artificial Intelligence (GenAI) on Knowledge Seeking in Online Communities," *Journal of Management Information Systems* (42:2), Informa UK Limited, pp. 370–399.
Venkatesh, V., Thong, J. Y. L., and Xu, X. 2012. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *Management Information Systems Quarterly* (36:1), MIS Quarterly, pp. 157–178.

W3

AI-Generated Review Summaries and Sales Mode Selection in Dual-Channel Supply Chains

Jiayao Zhao, Lin Wang, Zhiyong Li, and Hongxin Liu (Communication University of China).

AI-generated review summaries (AIGS) are increasingly adopted by e-commerce platforms to reduce information overload and improve review efficiency. However, existing research mainly examines consumer responses to AIGS and pays limited attention to its strategic implications for firms. This study develops a game-theoretic model of a dual-channel supply chain in which a manufacturer sells through an offline direct channel and an online platform under either the wholesale mode or the agency mode. We compare equilibrium outcomes with and without AIGS adoption, where AIGS improves review informativeness but requires platform investment. The results show that the agency mode always induces a higher online retail price than the wholesale mode, while the offline price ranking depends on review precision and AI enhancement. In addition, AIGS always strengthens the manufacturer's preference for the agency mode, whereas the platform prefers the agency mode only when AI investment cost is sufficiently low. Thus, AIGS not only improves information efficiency, but also reshapes governance preferences in dual-channel supply chains. Our study extends the literature by endogenizing AI-based information design in selling-mode choice.

KEY WORDS

AI-generated review summaries, dual-channel supply chain, selling-mode choice, e-commerce platform, wholesale mode, agency mode

AI-Generated Review Summaries and Sales Mode Selection in Dual-Channel Supply Chains

Zhao, J. Y.[a] , Wang, L.[b] , Li, Z. Y.[c] and Liu, H. X.[d]
[a] School of Economics and Management, Communication University of China
E-mail: jiayaozhao@mails.cuc.edu.cn
[b] School of Economics and Management, Communication University of China
E-mail: carolinewang@cuc.edu.cn
[c] School of Economics and Management, Communication University of China
E-mail: zyli@cuc.edu.cn (corresponding author)
[d] School of Economics and Management, Communication University of China
E-mail: Liuhongxin@cuc.edu.cn (corresponding author)

Abstract

AI-generated review summaries (AIGS) are increasingly adopted by e-commerce platforms to reduce information overload and improve consumers' access to review information. However, existing research mainly examines consumer responses to AIGS and pays limited attention to its strategic implications for firms. This study develops a game-theoretic model of a dual-channel supply chain in which a manufacturer sells through an offline direct channel and an online platform under either the wholesale mode or the agency mode. We compare equilibrium outcomes with and without AIGS adoption. The results show that the agency mode always leads to a higher online retail price, while the offline price depends on review precision and AI enhancement. Moreover, AIGS strengthens the manufacturer's preference for the agency mode, while the platform prefers the agency mode only when AI investment cost is sufficiently low. Our study highlights the role of AIGS in reshaping governance preferences in dual-channel supply chains.

Keywords: AI-generated review summaries, dual-channel supply chain, selling-mode choice, e-commerce platform, wholesale mode, agency mode

Introduction

With the rapid development of generative AI (Gen-AI), e-commerce platforms are increasingly adopting AI-generated review summaries (AIGS) to reorganize large volumes of online reviews and present product information in a more concise and accessible form. This trend has already appeared in industry practice. For example, Amazon introduced AI-generated review highlights on product detail pages to summarize frequently mentioned product features and customer sentiments, while JD's Gen-AI shopping assistant provides review summaries to help consumers compare products more efficiently (shown in Figure 1). AIGS is becoming an important digital tool for reducing consumers' information-processing burden and improving shopping efficiency.

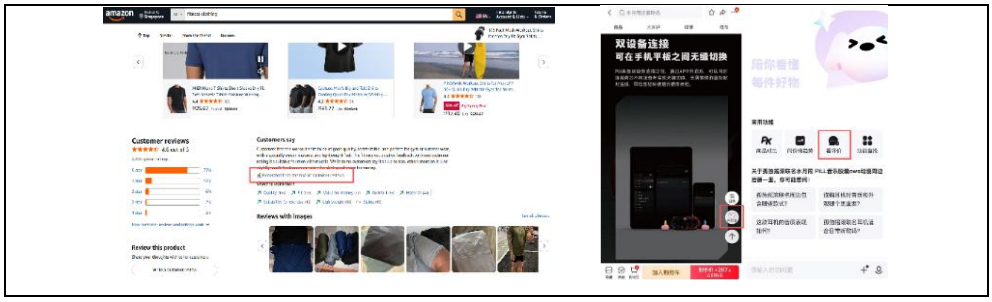


Figure 1. Amazon and JD Product Review with AIGS

The implications of AIGS go beyond consumer convenience. In a dual-channel supply chain, AIGS can change not only how quickly consumers process reviews, but also how accurately they interpret product information. Higher-quality summaries make review information more decision-relevant and easier to use, thereby affecting consumers' channel choices between the online and offline channels. These demand shifts, in turn, influence firms' pricing decisions, profit allocation, and the relative attractiveness of wholesale and agency selling arrangements. Therefore, AIGS should be viewed not merely as a consumer-facing feature, but as a strategic information-design tool in platform-based supply chains. In our model, this role is captured by the effective precision of review information. Existing studies have examined wholesale versus agency selling in dual-channel supply chains, the role of Gen-AI in e-commerce and operations management, and the economic effects of online reviews. However, these streams are largely studied in isolation. The dual-channel literature usually treats the information environment as exogenous, while the emerging AIGS literature mainly focuses on consumer response and platform information design. As a result, little is known about how AIGS affects equilibrium pricing and selling-mode choice in platform-based supply chains.

To address this gap, we develop a stylized dual-channel supply chain model with one manufacturer and one e-commerce platform. The manufacturer sells through an offline direct channel and an online channel operated through the platform. For the online channel, we consider both the wholesale mode and the agency mode. We study two scenarios, with and without AIGS adoption, where AIGS improves the effective quality of review information and eliminates consumers' review-reading cost, but requires platform investment.

We focus on the following research questions.

1. How does AIGS affect equilibrium prices under the wholesale and agency modes?
2. How does AIGS reshape the profits of the manufacturer and the platform under different selling arrangements?
3. How do AIGS enhancement efficiency and AIGS investment cost influence the selling-mode preferences of the manufacturer and the platform?

To answer these questions, we first derive the equilibrium outcomes in the benchmark case without AIGS under both selling modes. We then analyze how AIGS adoption changes prices and profits by improving review informativeness and reducing consumers' information-processing cost. Finally, we compare the two selling modes in the AIGS scenario and identify the conditions under which the manufacturer and the platform prefer different arrangements.

Literature Review

Our research is primarily related to three streams of literature: (1) wholesale versus agency selling in dual-channel supply chains, (2) Gen-AI in e-commerce and operations management, and (3) online reviews.

Wholesale versus agency selling in dual-channel supply chains

The wholesale mode and the agency mode differ mainly in the allocation of pricing power. Under the wholesale mode, the manufacturer sets the wholesale price and the platform sets the retail price. Under

the agency mode, the manufacturer directly controls online retail pricing and pays the platform a commission. The platform's preference between the two modes depends on how online sales interact with traditional channels: the agency mode is preferred when online sales cannibalize offline demand, whereas the wholesale mode is more attractive when online sales stimulate traditional demand (Abhishek et al. 2016).

More closely related to our study, Kwark et al. (2017) incorporated third-party information into the comparison between the two modes and found that different types of information may favor different channel arrangements. In particular, quality information tends to favor the wholesale mode, whereas fit information tends to favor the agency mode. However, this literature generally treats the information environment as given. It does not consider the possibility that the platform may actively reshape how review information is processed. In our study, this role is played by AI-generated review summaries (AIGS).

AI-generated summaries (AIGS) in e-commerce and operations management

Recent studies have begun to examine how generative AI is changing digital commerce. One increasingly important application is the use of AI-generated review summaries (AIGS), which condense large volumes of reviews into concise summaries. More recent work has started to examine consumers' responses to AIGS in more detail. Deng et al. (2026) studied GenAI-powered review summaries in online travel platforms and found that such summaries can increase review intention by reducing information overload and encouraging users to reciprocate the benefit they receive. Luo et al. (2026) showed that when AIGS is inconsistent with aggregated ratings, user trust declines, although this effect can be moderated by summary valence and algorithmic explanation. While these studies improve our understanding of how consumers respond to AIGS, they say little about how AIGS affects firms' strategic decisions in supply chain settings.

Online reviews

Online reviews are an important source of product information in e-commerce. They help reduce information asymmetry and influence consumer decisions (Chevalier and Mayzlin 2006). At the same time, the large volume of review content often creates information overload, making it costly for consumers to identify useful signals (Jones et al. 2004). Jabr and Rahman (2022) showed that platforms can partly address this problem by selectively featuring top reviews.

In summary, although the existing literature has generated important insights into channel mode choice, generative AI applications, and online reviews, these research streams have largely evolved in parallel. The dual-channel supply chain literature typically treats the information environment as exogenous and pays limited attention to how emerging AI tools may reshape market interactions. Meanwhile, the literature on AI-generated summaries (AIGS) focuses primarily on consumer perceptions and behavioral responses, without sufficiently examining their strategic implications for upstream and downstream firms. In addition, the online review literature has extensively studied the informational role of consumer reviews, yet it seldom considers AI-generated summaries as a novel mechanism for processing and presenting review information. Although recent studies, such as Deng et al. (2026) and Luo et al. (2026), offer valuable evidence on consumer responses to AIGS, they do not address how AIGS influences the strategic interaction between manufacturers and platforms. By integrating these strands of research, our study develops a unified framework to investigate how AIGS reshapes equilibrium pricing decisions and selling-mode preferences in a dual-channel supply chain.

Model

Consider a dual-channel supply chain consisting of one manufacturer and one e-commerce platform. The manufacturer sells a product through an offline direct channel and an online channel operated through the platform. Following common practice in e-commerce, we consider two online selling modes: the wholesale mode and the agency mode. Under the wholesale mode, the manufacturer first sets the wholesale price w for the online channel and the offline retail price p_d. After observing these decisions, the platform determines the online retail price p_r. Under the agency mode, the manufacturer

sells directly to consumers through the platform and pays the platform a commission rate α, where $0 < \alpha < 1$. In this case, the manufacturer simultaneously determines the online retail price p_r and the offline retail price p_d.

To capture consumers' channel preferences, we normalize the valuation coefficient of the online channel to 1 and denote that of the offline channel by θ, where $0 < \theta < 1$. Thus, relative to the online channel, consumers derive a lower basic valuation from the offline channel. Let $q_0 \in (0,1)$ denote the positive value information reflected by online reviews. To focus on the role of AIGS in improving review informativeness, we restrict attention to the case of favorable review information only. Let η denote the precision of review information. In our model, AIGS quality is captured parsimoniously by the effective precision term $\beta\eta$, where $\beta > 1$ measures the enhancement effect of AIGS on review informativeness. A higher value of $\beta\eta$ means that the generated summary provides more decision-relevant review information to consumers, which changes channel choice thresholds and, in turn, affects firms' pricing and governance decisions. In addition, the platform incurs an AI investment cost $k\beta^2$, where $k > 0$.

Accordingly, under AIGS adoption, the net utility of a consumer purchasing through the online channel is given by

$$U_r = (1 - \beta\eta)q + \beta\eta q_0 - p_r.$$

whereas the net utility of purchasing through the offline channel is

$$U_d = (1 - \beta\eta)\theta q + \beta\eta q_0 - p_d.$$

When $U_r = 0$, the consumer is indifferent between purchasing through the online channel and not purchasing, yielding the threshold $q_r = \frac{-p_r + \beta\eta q_0}{-1 + \beta\eta}$. When $U_d = 0$, the consumer is indifferent between purchasing through the offline channel and not purchasing, yielding the threshold $q_d = \frac{-p_d + \beta\eta q_0}{(-1 + \beta\eta)\theta}$. When $U_r = U_d$, the consumer is indifferent between the online and offline channels, yielding the threshold $q_{rd} = \frac{p_r - p_d}{(-1 + \beta\eta)(-1 + \theta)}$. To ensure the coexistence of the online and offline channels, the following conditions must hold: $w \le p_d < \beta\eta q_0 + \theta - \beta\eta\theta$ and $\frac{p_d - q_0\beta\eta + q_0\beta\eta\theta}{\theta} < p_r < 1 + p_d - \beta\eta - \theta + \beta\eta\theta$.

Accordingly, the demand functions of the two channels can be written as:

$$D_r = 1 + \frac{p_d - p_r}{(-1 + \beta\eta)(-1 + \theta)},$$
$$D_d = \frac{-p_d + \eta q_0 + (p_r - \eta q_0)\theta}{(-1 + \beta\eta)(-1 + \theta)\theta}.$$

Benchmark Case: Without AIGS

Wholesale Mode

Under the wholesale mode, the manufacturer determines the wholesale price $w$and the offline retail price p_d^W, while the platform subsequently sets the online retail price p_r^W. The manufacturer's and the platform's profit functions are given by

$$\pi_m^W = wD_r^W + p_d^W D_d^W,$$
$$\pi_r^W = (p_r^W - w)D_r^W.$$

By solving the platform's pricing problem first and then substituting the optimal response into the manufacturer's problem, we obtain the following equilibrium outcomes.

Lemma 1 (Wholesale mode without AIGS): Without AIGS, the equilibrium decisions under the wholesale mode are given by $w^* = \frac{1}{2}(\eta(q_0 - \theta) + \theta - r)$, $p_d^{W*} = \frac{1}{2}(\eta(q_0 - \theta) + \theta - r)$, $p_r^{W*} =$

$\frac{1}{2}(1+\eta(q_0-1)-r)$. The corresponding equilibrium profits are $\pi_m^{W^*}=-\frac{(r-q_0\eta+(-1+\eta)\theta)^2}{4(-1+\eta)\theta}$, $\pi_r^{W^*}=\frac{1}{4}(-1+\eta)(-1+\theta)$.

Agency Mode

Under the agency mode, the manufacturer directly sells through the platform and pays the platform a commission rate α. The manufacturer simultaneously determines the online retail price p_r^A and the offline retail price p_d^A. The profit functions are

$$\pi_m^A=(1-\alpha)p_r^A D_r^A+p_d^A D_d^A,$$
$$\pi_r^A=\alpha p_r^A D_r^A.$$

Solving the manufacturer's optimization problem yields the following result.

Lemma 2 (Agency mode without AIGS): The equilibrium decisions under the agency mode are given by $p_r^{A^*}=\frac{(\theta-1)(2+r(\alpha-2)-2\alpha+\eta(-2+2q_0+2\alpha-q_0\alpha))}{\theta(\alpha-2)^2+4(\alpha-1)}$, $p_d^{A^*}=-\frac{(\theta-1)(\alpha-1)(2\eta q_0-2r-\theta(\alpha-2)+\eta\theta(\alpha-2))}{\theta(\alpha-2)^2+4(\alpha-1)}$. The corresponding profits are $\pi_m^{A^*}=\frac{(\theta-1)(\alpha-1)\Omega_M}{(\eta-1)\theta(\theta(\alpha-2)^2+4(\alpha-1))}$, $\pi_r^{A^*}=\frac{(\theta-1)\alpha\,\Omega_{R1}\Omega_{R2}}{(\eta-1)(\theta(\alpha-2)^2+4(\alpha-1))^2}$, where $\Omega_M=r^2+\theta-\eta\left[q_0\left(2r+\theta(\alpha-2)\right)+\theta(2+r(\alpha-2)-2\alpha)\right]+r\theta(\alpha-2)-\theta\alpha+\eta^2[q_0^2+\theta+\theta(\alpha-2)q_0-\theta\alpha]$, $\Omega_{R1}=-2+2r+\eta(2+q_0(\alpha-2)-2\alpha)+2\alpha-r$, $\Omega_{R2}=-2-\theta(\alpha-2)+2\alpha-r\alpha+\eta(2+\theta(\alpha-2)+(q_0-2)\alpha)$.

Compare the equilibrium prices in Lemma 1 and Lemma 2, we have the following proposition 1.

Proposition 1: (i) $p_r^{A^*}>p_r^{W^*}$.

(ii) There exists a threshold $\eta_0^N(\theta)=\frac{2-2\alpha+r\alpha-2\theta+\alpha\theta}{2-2\alpha+q_0\alpha-2\theta+\alpha\theta}$, such that $p_d^{W^*}>p_d^{A^*}$ if $\eta<\eta_0^N(\theta)$, and $p_d^{W^*}<p_d^{A^*}$ if $\eta>\eta_0^N(\theta)$.

Without AIGS, the agency mode always leads to a higher online retail price than the wholesale mode because the manufacturer retains direct pricing authority in the online channel. By contrast, the offline price ranking depends on review precision, since better review information changes consumers' channel preferences and intensifies cross-channel competition.

AIGS Adoption Case

In this case, AIGS improves the effective precision of review information from η to $\beta\eta$, eliminates consumers' review-reading cost, and imposes an AI investment cost $k\beta^2$ on the platform. Relative to the benchmark case, AIGS affects market outcomes by changing the demand structure and the payoff functions.

Wholesale Mode With AIGS

Under this mode with AIGS, the manufacturer's and the platform's profit functions are given by

$$\bar{\pi}_m^W=\bar{w}\bar{D}_r^W+\bar{p}_d^W\bar{D}_d^W,$$
$$\bar{\pi}_r^W=(\bar{p}_r^W-\bar{w})\bar{D}_r^W-k\beta^2.$$

Using backward induction, we obtain the following equilibrium result.

Lemma 3 (Wholesale mode with AIGS): When the platform adopts AIGS, the equilibrium decisions under the wholesale mode are given by $\bar{w}^* = \frac{1}{2}(\beta\eta(q_0-\theta)+\theta)$, $\bar{p}_d^{W^*} = \frac{1}{2}(\beta\eta(q_0-\theta)+\theta)$, $\bar{p}_r^{W^*} = \frac{1}{2}(1+(-1+q_0)\beta\eta)$. The corresponding equilibrium profits are $\bar{\pi}_m^{W^*} = -\frac{(q_0\beta\eta+\theta-\beta\eta\theta)^2}{4(-1+\beta\eta)\theta}$, $\bar{\pi}_r^{W^*} = -k\beta^2 + \frac{1}{4}(-1+\beta\eta)(-1+\theta)$.

Agency Mode With AIGS

Under the agency mode with AIGS, the manufacturer's and the platform's profit functions are given by

$$\bar{\pi}_m^A = (1-\alpha)\bar{p}_r^A\bar{D}_r^A + \bar{p}_d^A\bar{D}_d^A,$$
$$\bar{\pi}_r^A = \alpha\bar{p}_r^A\bar{D}_r^A - k\beta^2.$$

To ensure the coexistence of both channels and the feasibility of equilibrium outcomes, we impose the relevant parameter restrictions and solve the constrained optimization problem using the Lagrangian method and KKT conditions. We then have the following result.

Lemma 4 (Agency mode with AIGS): When $0<\alpha<1$, $0<q_0<1$, $\beta>1$, $0<\theta<\frac{-2+2\alpha}{-2+\alpha}$, and $0<\eta\le\eta_1(\theta)$, the equilibrium decisions under the agency mode are given by $\bar{p}_r^{A^*} = -\frac{(2(-1+\alpha)+(2+q_0(-2+\alpha)-2\alpha)\beta\eta)(-1+\theta)}{4(-1+\alpha)+(-2+\alpha)^2\theta}$, $\bar{p}_d^{A^*} = -\frac{(-1+\alpha)(-1+\theta)(2q_0\beta\eta+(-2+\alpha)(-1+\beta\eta)\theta)}{4(-1+\alpha)+(-2+\alpha)^2\theta}$, where $\eta_1(\theta) = \frac{2-2\alpha-2\theta+\alpha\theta}{2\beta-2\alpha\beta+q_0\alpha\beta-2\beta\theta+\alpha\beta\theta}$. The corresponding equilibrium profits are given by $\bar{\pi}_m^{A^*} = \frac{(-1+\alpha)(-1+\theta)(q_0^2\beta^2\eta^2+q_0(-2+\alpha)\beta\eta(-1+\beta\eta)\theta-(-1+\alpha)(-1+\beta\eta)^2\theta)}{(-1+\beta\eta)\theta(-4+4\alpha+4\theta-4\alpha\theta+\alpha^2\theta)}$, $\bar{\pi}_r^{A^*} = \frac{\alpha(-2-2(-1+q_0)\beta\eta+\alpha(2+(-2+q_0)\beta\eta))(-1+\theta)(-2(-1+\beta\eta)(-1+\theta)+\alpha(2-\theta+\beta\eta(-2+q_0+\theta)))}{(-1+\beta\eta)(-4+4\alpha+4\theta-4\alpha\theta+\alpha^2\theta)^2}$.

Compare the equilibrium prices in Lemma 3 and Lemma 4, we have the following proposition 2.

Proposition 2: (i) $\bar{p}_r^{A^*} > \bar{p}_r^{W^*}$. (ii) Let $\eta_1(\theta) = \frac{2-2\alpha-2\theta+\alpha\theta}{2\beta-2\alpha\beta+q_0\alpha\beta-2\beta\theta+\alpha\beta\theta}$. Then, $\bar{p}_d^{W^*} > \bar{p}_d^{A^*}$, if $\eta < \eta_1(\theta)$, and $\bar{p}_d^{W^*} < \bar{p}_d^{A^*}$, if $\eta > \eta_1(\theta)$.

With AIGS, the agency mode still leads to a higher online retail price than the wholesale mode, because the manufacturer retains direct control over online pricing. By contrast, the offline price ranking depends on a new threshold shaped by review precision and AI enhancement. Therefore, AIGS preserves the online pricing order while reshaping the offline price comparison through its effect on information efficiency and channel substitution.

Revenue Mode Preference

We next compare the equilibrium profits under the wholesale mode and the agency mode in the AIGS scenario, and characterize the mode preferences of the manufacturer and the platform. The profit difference between the two modes can be written as:

$$\bar{\pi}_m^A - \bar{\pi}_m^W = \frac{\left[-2(-1+\beta\eta)(-1+\theta)+\alpha\left(2-\theta+\beta\eta(-2+q_0+\theta)\right)\right]^2}{4(-1+\beta\eta)(4(-1+\alpha)+(-2+\alpha)^2\theta)},$$
$$\bar{\pi}_r^A - \bar{\pi}_r^W = k\beta^2 - \frac{1}{4}(-1+\beta\eta)(-1+\theta) + \frac{\alpha(A+C\beta\eta)(\theta-1)(B+D\beta\eta)}{(\beta\eta-1)(-4+4\alpha+4\theta-4\alpha\theta+\alpha^2\theta)^2}.$$

Based on the above analysis, we have the following propositions 3 and 4.

Proposition 3: Under AIGS adoption, the manufacturer always prefers the agency mode, i.e., $\bar{\pi}_m^A > \bar{\pi}_m^W$.

From Proposition 3, we find that AIGS strengthens the manufacturer's incentive to choose the agency mode. The reason is that the agency mode gives the manufacturer direct control over retail pricing, allowing it to better appropriate the value generated by improved review informativeness. As AIGS enhances the informational role of reviews, pricing authority becomes more valuable, which makes the agency mode consistently more attractive to the manufacturer.

Proposition 4: Denoted $k^*(\beta) = \frac{\theta-1}{\beta^2}\left[\frac{1}{4}(\beta\eta - 1) - \frac{\alpha(\theta-1)(A+C\beta\eta)(B+D\beta\eta)}{(\beta\eta-1)(-4+4\alpha+4\theta-4\alpha\theta+\alpha^2\theta)^2}\right]$, where $A = -2 + 2\alpha,\ B = 2\theta - \theta\alpha - 2 + 2\alpha,\ C = 2 - 2q_0 - 2\alpha + q_0\alpha,\ D = q_0\alpha + \theta\alpha - 2\alpha - 2\theta + 2$. With AIGS adoption, $\bar{\pi}_r^A > \bar{\pi}_r^W$ if $k < k^*(\beta)$, and $\bar{\pi}_r^A < \bar{\pi}_r^W$ if $k > k^*(\beta)$.

From Proposition 4, we find that the platform's mode preference depends on the trade-off between the benefits of AIGS and its investment cost. When the AI investment cost is sufficiently low, the platform can benefit more from the efficiency gains brought by AIGS under the agency mode. However, when the investment cost becomes high, such gains are no longer sufficient to compensate for the cost burden, making the wholesale mode more attractive to the platform. Therefore, the platform's preference is conditional rather than fixed under AIGS adoption.

Based on Propositions 3 and 4, we further summarize the joint preference structure of the manufacturer and the platform under AIGS adoption in proposition 5.

Proposition 5: (i) if $k < k^*(\beta)$, then $\bar{\pi}_m^A > \bar{\pi}_m^W$ and $\bar{\pi}_r^A > \bar{\pi}_r^W$, that is, both the manufacturer and the platform prefer the agency mode;

(ii) if $k > k^*(\beta)$, then $\bar{\pi}_m^A > \bar{\pi}_m^W$ and $\bar{\pi}_r^A < \bar{\pi}_r^W$, that is, the manufacturer prefers the agency mode, whereas the platform prefers the wholesale mode.

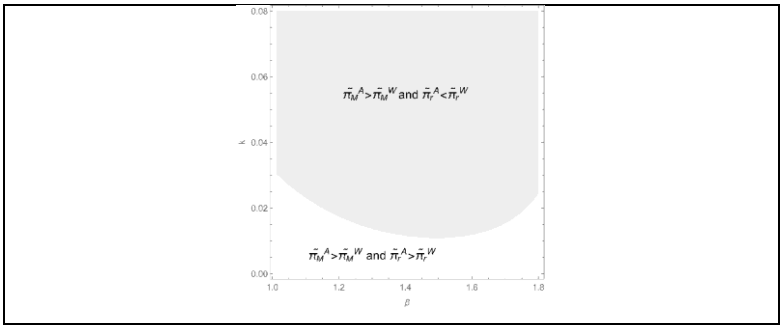


Figure 2. Profit Preference Regions of Agency and Wholesale Modes ($q_0 = 0.6,\ \eta = 0.5,\ \alpha = 0.3,\ \theta = 0.4$).

Figure 2 provides a numerical illustration of Propositions 3–5 under a representative parameter setting ($q_0 = 0.6,\ \eta = 0.5,\ \alpha = 0.3,\ \theta = 0.4$). As shown in Figure 2, AIGS adoption generates two distinct preference regions in the (β, k) plane. Better review summaries increase the strategic importance of pricing control, which strengthens the manufacturer's preference for the agency mode. By contrast, the platform evaluates AIGS adoption from a cost–benefit perspective. When AI investment cost is low, both firms prefer the agency mode. When the investment cost is high, the platform can no longer fully recover its technology expenditure, and a divergence in mode preferences emerges. This numerical example illustrates how AIGS reshapes channel governance through the interaction of information enhancement and technology cost.

Conclusion

This study investigates how AI-generated review summaries (AIGS) affect equilibrium pricing and selling-mode selection in a dual-channel supply chain consisting of one manufacturer and one e-commerce platform. By modeling AIGS as an information-enhancing technology that improves review informativeness while requiring platform investment, we show that AIGS reshapes both pricing outcomes and governance preferences across wholesale and agency arrangements. In particular, the agency mode always yields a higher online retail price than the wholesale mode, both with and without AIGS, whereas the offline price comparison depends on review informativeness and, under AIGS, also on the AI enhancement effect. These findings suggest that AIGS does not simply improve information efficiency, but also changes the strategic interaction between channels.

Our analysis further reveals that AIGS has asymmetric implications for the manufacturer and the platform. For the manufacturer, AIGS strengthens the attractiveness of the agency mode because improved review informativeness increases the value of retaining direct pricing control. For the platform, however, the preference between agency and wholesale depends on whether the efficiency gains from AIGS are sufficient to offset the corresponding AI investment cost. As a result, AIGS generates two distinct preference regions: when the investment cost is low, both firms prefer the agency mode; when the investment cost is high, the manufacturer still prefers the agency mode, while the platform switches to the wholesale mode. Therefore, AIGS may either align or diverge channel members' governance preferences.

This study contributes to the literature in two ways. First, it extends the dual-channel supply chain literature by endogenizing the review information environment and showing how an AI-based information tool can affect equilibrium prices and selling-mode choice. Second, it contributes to the emerging research on Gen-AI in e-commerce by moving beyond consumer response and highlighting the strategic implications of AIGS for upstream and downstream firms. From a managerial perspective, our findings suggest that manufacturers have stronger incentives to support agency arrangements after AIGS adoption, while platforms should evaluate AI deployment not only in terms of consumer convenience, but also in terms of governance consequences and cost recovery. Future research may consider heterogeneous consumer trust in AIGS, endogenous commission design, and more competitive supply-chain structures with multiple manufacturers or platforms.

Acknowledgements

This research is supported by research grant from the National Natural Science Foundation of China (No. 72472143) and the Fundamental Research Funds for the Central Universities (No. CUCZL2508).

References

Abhishek, V., Jerath, K., and Zhang, Z. J. 2016. "Agency Selling or Reselling? Channel Structures in Electronic Retailing," *Management Science* (62:8), pp. 2259–2280.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345–354.

Deng, L., Li, C., and Ye, Q. 2026. "Cool or Poor: Assessing the Effectiveness of Genai-Powered Review Summary Feature on the Ota Platform," *Tourism Management* (114).

Jabr, W., and Rahman, M. S. 2022. "Online Reviews and Information Overload: The Role of Selective, Parsimonious, and Concordant Top Reviews," *MIS Quarterly* (46:3), pp. 1517–1550.

Jones, Q., Ravid, G., and Rafaeli, S. 2004. "Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* (15:2), pp. 194–210.

Kwark, Y., Chen, J., and Raghunathan, S. 2017. "Platform or Wholesale? A Strategic Tool for Online Retailers to Benefit from Third-Party Information," *MIS Quarterly* (41:3), pp. 763–A717.

Luo, L., Liu, L., and Zheng, Y. 2026. "When Ai Summary Is Inconsistent with Aggregated Rating: The Role of Trust, Valence, and Algorithmic Explanation in Consumer Attitudes," *Telematics and Informatics* (104), p. 102361.

W3

Conversational AI in High-Involvement Purchases: Chatbot-Assisted EV Buying and Channel Choice in Germany.

Tianhao Li (University of Göttingen)

*교신저자

As AI transforms high-value product marketing, its influence on complex consumer decision-making remains underexplored. This study investigates the impact of conversational AI-based sales chatbots on electric vehicle (EV) purchase decisions. Using a quantitative survey and structural equation modeling, the research examines how AI-assisted services (AAS) help consumers navigate information overload and influence channel choice (CC), which in turn affects electric vehicle purchase decisions (EPD). In addition, the study explores the mediating role of CC in the relationship between AAS and EPD, as well as the moderating effects of symbolic purchase motives (SPM) and perceived manipulative intent (PMI) on the relationship between AAS and CC.

KEY WORDS

AI Sales Chatbot, EV Purchase Decision, Channel Choice, Symbolic Purchase Motives, Perceived Manipulative Intent

Conversational AI in High-Involvement Purchases: Chatbot-Assisted EV Buying and Channel Choice in Germany

Li, T.
Faculty of Business and Economics, University of Göttingen
E-mail: tianhao.li@stud.uni-goettingen.de

Abstract

As AI transforms high-value product marketing, its influence on complex consumer decision-making remains underexplored. This study investigates the impact of conversational AI-based sales chatbots on electric vehicle (EV) purchase decisions. Using a quantitative survey and structural equation modeling, the research examines how AI-assisted services (AAS) help consumers navigate information overload and influence channel choice (CC), which in turn affects electric vehicle purchase decisions (EPD). This study examines whether CC mediates the relationship between AAS and EPD, as well as the moderating effects of symbolic purchase motives (SPM) and perceived manipulative intent (PMI) on the relationship between AAS and CC.

Keywords: AI-based sales Chatbot, Electric Vehicles Purchase Decision, Channel Choice, Symbolic Purchase Motives, Perceived Manipulative Intent

Introduction

High-value products, such as electric vehicles (EVs), involve high cost, infrequent purchase, and complex decision-making processes (Nagle et al., 2023). While prior research has examined EV adoption and digital sales tools separately (e.g., Frank et al., 2024; Nagy et al., 2021), limited attention has been paid to how AAS, such as conversational chatbots, influence consumer decision-making. Unlike prior studies that primarily examine AI adoption in low-involvement online shopping contexts, this study extends the literature to high-involvement durable purchases such as electric vehicles.

In this study, AAS are conceptualized as a sales chatbot that provides personalized recommendations and purchase support through natural language interaction. By reducing information uncertainty and enhancing perceived convenience during high-involvement purchase decisions, AI sales chatbots may increase consumers' preference for online channels. In the context of EV purchases, online channels provide efficient information search and time-saving convenience, whereas offline channels offer trust-building opportunities, face-to-face consultation, and test-drive experiences. CC is particularly important in high-value purchases, as online and offline channels offer distinct advantages in terms of efficiency and experiential value (Wolf & Steul-Fischer, 2023).

However, the effectiveness of AAS may depend on important boundary conditions. Consumers with strong SPM may resist AI recommendations that conflict with identity or status goals (Noppers et al., 2014; Belk, 1988). In addition, PMI may reduce trust in AI systems and weaken their influence. Taken together, these arguments suggest that channel choice represents a critical mechanism through which AI-assisted sales influence final purchase decisions.

Research Model

Research Framework

This study examines how AAS influence EPD through CC, while considering the moderating roles of SPM and PMI. More importantly, it positions CC as a key decision mechanism and identifies SPM and PMI as important boundary conditions in high-involvement purchase decisions.

Research Hypotheses

Based on the integrated framework, the following hypotheses will be tested to explore the relationships between the identified variables:

H1: AAS is expected to positively influence consumers' preference for online CC.

H2: CC is hypothesized to positively influence EPD.

H3: AAS is anticipated to be positively associated with EPD.

H4: CC is expected to mediate the relationship between AAS and EPD.

H5: SPM will negatively moderate the relationship between AAS and CC.

H6: PMI will negatively moderate the effect of AAS on CC.

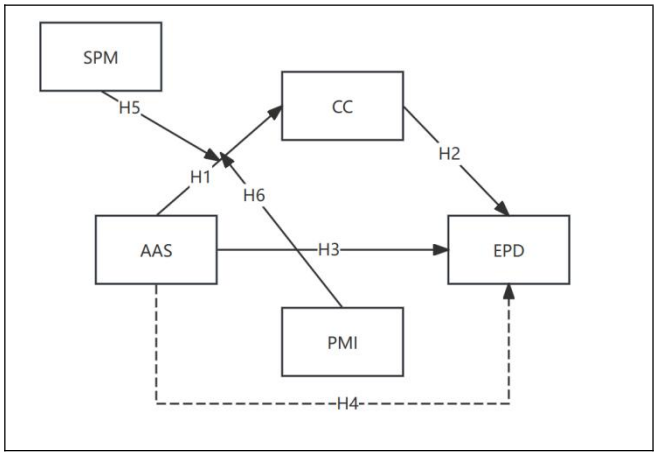


Figure 1. Research Model

Research Methodology

Research Design and Data Collection

This study will employ a quantitative survey design to test the proposed theoretical model. Data will be collected from consumers residing in Germany who intend to purchase a vehicle within the next two years and have considered EV alternatives.

Participants will first be presented with a standardized EV purchase scenario involving an AI sales chatbot that provides consultation, recommendations, and channel guidance. They will then complete measures of symbolic motives, PMI, CC, and EPD. Data will be collected via online research platforms to ensure demographic diversity.

Variable Measurement

The study will employ established scales adapted to the context of EV purchases, using a 7-point Likert scale (1 = "Strongly Disagree", 7 = "Strongly Agree").

AAS will be measured using AI objectivity scales adapted from Pan et al. (2022), focusing on the perceived accuracy and neutrality of chatbot recommendations.

CC will be adapted from Kushwaha and Shankar (2013) to assess consumers' relative preference for online platforms versus offline dealerships.

EPD will be measured through purchase intention items adapted from Dodds et al. (1991). SPM will be based on Noppers et al. (2014), capturing identity-signaling and status-related motives for choosing an EV.

PMI will be adapted from Campbell (1995), measuring the extent to which consumers perceive the AI's recommendations as being influenced by the seller's commercial interests.

Control variables will include age, income, educational attainment, and prior knowledge of EVs.

Data Analysis Plan

Data will be analyzed using SPSS and AMOS. Reliability will be assessed using Cronbach's alpha, and CFA will be conducted in AMOS to validate the measurement model. The proposed moderated mediation effects will then be tested using PROCESS Model 7 in SPSS with 5,000 bootstrap samples.

References

Campbell, M. C. 1995. "When Attention-Getting Advertising Tactics Elicit Consumer Skepticism: The Role of Intent and Language," Journal of Consumer Psychology (4:3), pp. 225-254.

Dodds, W. B., Monroe, K. B., and Grewal, D. 1991. "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," Journal of Marketing Research (28:3), pp. 307-319.

Frank, D.-A., Chrysochou, P., Mitkidis, P., Otterbring, T., and Ariely, D. 2024. " Navigating Uncertainty: Exploring Consumer Acceptance of Artificial Intelligence under Self-Threats and High-Stakes Decisions," Technology in Society (79), Article 102732.

Kushwaha, T., and Shankar, V. 2013. " Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics," Journal of Marketing (77:4), pp. 67-85.

Nagle, T. T., Müller, G., and Gruyaert, E. 2023. The Strategy and Tactics of Pricing: A Guide to Growing More Profitably (7th ed.). New York: Routledge.

Nagy, S., and Hajdu, N. 2021. "Consumer Acceptance of the Use of Artificial Intelligence in Online Shopping: Evidence from Hungary," Amfiteatru Economic (23:56), pp. 155–173.

Noppers, E. H., Keizer, K., Bockarjova, M., and Steg, L. 2014. " The Adoption of Sustainable Innovations: Driven by Symbolic and Environmental Motives," Global Environmental Change (25), pp. 52-62.

Pan, Y., Froese, F. J., Liu, N., Hu, Y., and Ye, M. 2022. "The Adoption of Artificial Intelligence in Employee Recruitment: The Influence of Contextual Factors," International Journal of Human Resource Management (33:6), pp. 1125-1147.

Wolf, L., and Steul-Fischer, M. 2023. "Factors of Customers' Channel Choice in an Omnichannel Environment: A Systematic Literature Review," Management Review Quarterly (73), pp. 1579–1630.

W3

From Features to Semantics: Content Reconstruction with LLMs for Recommendation.

Dahee Kim, Seonu Park (Kyung Hee University),
Qinglong Li (Hansung University) and Sung-Byung Yang (Kyung Hee University)

Information overload in e-commerce hinders user decision-making, underscoring the importance of recommender systems. Rating prediction is a core task, but existing methods struggle to semantically integrate heterogeneous data and handle missing values without information loss. This study proposes an LLM-based Content Reconstruction for Recommendation (LCRR) framework that serializes diverse user and item features into natural language, enabling unified semantic representations. Unlike conventional approaches, LCRR treats missing data as meaningful signals rather than removing or imputing them, preserving underlying semantics. The framework enhances representation learning by capturing both contextual and feature-level interactions while flexibly accommodating diverse data types and missing values. Experimental results show consistent improvements in RMSE and MAE across multiple datasets, demonstrating robustness under different sparsity and missing conditions. Overall, LCRR introduces a novel paradigm for leveraging LLMs to integrate heterogeneous and incomplete data, improving recommendation accuracy and applicability in real-world scenarios.

KEY WORDS

Recommender System, Large Language Model, Serialization, Tabular Data

W3

Explaining E-Commerce Customer Engagement via Two-Part Modeling and LLM-Based Signal Analysis

Jiho Kim, Sangmin Park, Yooshin Kwon (Chung-Ang Univ.), Heongi Kim and Hyeonjae Cheon (GS Retail)

Online customer reviews differ substantially in the engagement they receive, with only a small proportion attracting helpful votes. This study examines review engagement by integrating predictive modeling and interpretable signal analysis. Drawing on signaling theory, review characteristics are conceptualized as observable cues of usefulness and informativeness. Linguistic features, such as length, readability, sentiment, and emphasis, are combined with topic-based features from two topic-modeling approaches. Given the zero-inflated distribution of helpful-vote counts, a two-part modeling approach is adopted to estimate vote occurrence and vote count conditional on engagement. Preliminary results indicate that a gradient-boosting model performs best, particularly in ranking potentially engaging reviews rather than predicting exact vote counts. Feature interpretation suggests that informational richness is more closely associated with engagement than sentiment intensity alone. An exploratory LLM-based analysis further examines whether generated reviews reproduce engagement-related signals found in human-written reviews.

KEY WORDS

Customer engagement, online customer reviews, topic modeling, explainable artificial intelligence, large language models

Explaining E-Commerce Customer Engagement via Two-Part Modeling and LLM-Based Signal Analysis

Jiho Kim[a,], Sangmin Park[a], Yooshin Kwon[a], Heongi Kim[b], and Hyeonjae Cheon[b]*
[a] Software Education Institute, Chung-Ang University
E-mail: jihonav@cau.ac.kr, sanggoe0509@cau.ac.kr, kwon0715@cau.ac.kr
[b] AI Data Division, GS Retail Co., Ltd.
E-mail: kim.hk@gsretail.com, slash@korea.ac.kr

Abstract

Online customer reviews differ substantially in the engagement they receive, with only a small proportion attracting helpful votes. This study examines review engagement by *integrating predictive modeling and interpretable signal analysis. Drawing on signaling theory, review characteristics are conceptualized as observable cues of usefulness and informativeness. Linguistic features, such as length, readability, sentiment, and emphasis, are combined with topic-based features from two topic-modeling approaches. Given the zero-inflated distribution of helpful-vote counts, a two-part modeling approach is adopted to estimate vote occurrence and vote count conditional on engagement. Preliminary results indicate that a gradient-boosting model performs best, particularly in ranking potentially engaging reviews rather than predicting exact vote counts. Feature interpretation suggests that informational richness is more closely associated with engagement than sentiment intensity alone. An exploratory LLM-based analysis further examines whether generated reviews reproduce engagement-related signals found in human-written reviews.*

Keywords: Customer engagement, online customer reviews, topic modeling, explainable artificial intelligence, large language models

Introduction

Online customer reviews have become a central source of information in electronic commerce, influencing consumers' purchase decisions and firms' competitive positioning. However, reviews differ substantially in their visibility and influence. While a small number of reviews accumulate many usefulness votes and gain prominent placement, most receive little or no attention (Mudambi and Schuff, 2010). Identifying the characteristics that distinguish engaging reviews from less visible ones is therefore important for both research and practice.

Prior studies have examined review usefulness using various textual, reviewer-related, and product-related characteristics (Krishnamoorthy, 2015; Yang et al., 2021). These studies suggest that observable review characteristics, such as length, clarity, sentiment, and topical content, may provide useful cues for assessing the perceived value of a review. From this perspective, review engagement can be understood not only as an outcome to be predicted, but also as a phenomenon that can be interpreted through the linguistic and topical signals embedded in review text.

As user-generated content continues to expand, platforms face increasing pressure to identify high-value reviews and reduce information overload (Park and Lee, 2008). Building on this stream of research, this study aims to predict review engagement using a two-part modeling approach suited to highly skewed usefulness vote distributions and to interpret the linguistic and topical characteristics associated with engagement. As an exploratory extension, this study further examines whether a large language model can reproduce engagement-related signals observed in human-written reviews.

Method

Data and features

This study uses 3,000 Amazon product reviews containing review text, summaries, product descriptions, and product related metadata. The outcome variable is the number of usefulness votes. As shown in Table 1, the vote distribution is highly zero inflated and long tailed, which guided the modeling strategy.

Table 1. Review vote distribution

Group	No votes	More than five votes	One to five votes
Proportion	approximately 86%	approximately 3%	approximately 11%

Reviews are represented using linguistic and topical features, including stylistic indicators and topic representations derived from latent Dirichlet allocation and BERTopic with Sentence BERT embeddings (Grootendorst, 2022). These features capture both the style and content of each review.

Two-Part Modeling Approach

Since usefulness votes are highly zero-inflated, we adopt a two-part hurdle formulation. The first component estimates vote occurrence, and the second estimates vote magnitude conditional on receiving at least one vote. Several gradient-boosting algorithms were compared for the second component. Model evaluation considers both prediction error and ranking quality, reflecting the practical goal of surfacing potentially useful reviews (Olmedilla et al., 2022).

Results

Model performance

Table 2 presents the comparative performance of the candidate models. CatBoost achieved the lowest prediction error and was selected for subsequent analysis. Although overall explanatory power remained modest, ranking performance was more promising, suggesting that the models are better suited to review prioritization than to exact vote-count prediction.

Table 2. Comparative model performance

Model	**RMSE**	**Ranking score**	**Hit rate (Top 10%)**
CatBoost	2.30	0.37	0.37
LightGBM	2.32	0.37	0.40
XGBoost	2.39	0.37	0.40

Features associated with engagement

As summarized in Table 3, informational richness emerged as the most consistent signal of engagement. Longer and more detailed reviews were associated with higher engagement, whereas positive sentiment showed a weaker association. This approach aligns with prior work combining feature engineering and explainable methods for review usefulness analysis (Kim et al., 2023; Lundberg and Lee, 2017).

Table 3. Features associated with engagement

Feature	**Association**	**Remark**
Review length	Positive	Longer reviews engage more
Number of sentences	Positive	Consistent with length
Positive sentiment	Weakly positive	Smaller effect than length

Exploratory Analysis of LLM-Generated Reviews

As an exploratory extension, a large language model, Llama-3.3-70B, was prompted to generate reviews resembling highly engaging examples. As shown in Table 4, generated reviews showed stronger positive sentiment but were shorter and less detailed than real high engagement reviews. This result further supports the importance of informational depth in review engagement.

Table 4. Comparison of real high-engagement and LLM-generated reviews

Characteristic	Real reviews	Generated reviews
Length	Longer	Shorter
Sentence count	Higher	Lower
Positive sentiment	Moderate	Stronger

Conclusion

This study provides preliminary evidence that informational depth is strongly associated with review engagement. Across model interpretation, qualitative text review, and LLM based analysis, longer and more detailed reviews showed a stronger relationship with engagement than sentiment alone. These findings suggest that ecommerce platforms may benefit from prioritizing reviews that provide specific product use information.

Several limitations remain. Review engagement may be influenced by display position, timing, product popularity, and exposure duration, which were not fully captured in this study. The small and category limited dataset also constrains generalizability. Future research should account more explicitly for these factors and test the findings across broader product categories and languages. Overall, this study suggests that review engagement can be examined as both a prediction task and an interpretable signal analysis problem.

Acknowledgements

This study was supported by GS Retail Co. Ltd.

References

Grootendorst, M. 2022. “BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure,” arXiv preprint arXiv:2203.05794.

Kim, J., Lee, H., and Lee, H. 2023. “Mining the Determinants of Review Helpfulness: A Novel Approach Using Intelligent Feature Engineering and Explainable AI,” Data Technologies and Applications (57:1), pp. 108-130.

Krishnamoorthy, S. 2015. “Linguistic Features for Review Helpfulness Prediction,” Expert Systems with Applications (42:7), pp. 3751-3759.

Lundberg, S. M., and Lee, S.-I. 2017. “A Unified Approach to Interpreting Model Predictions,” in Advances in Neural Information Processing Systems 30, pp. 4765-4774.

Mudambi, S. M., and Schuff, D. 2010. “What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com,” MIS Quarterly (34:1), pp. 185-200.

Olmedilla, M., Martínez-Torres, M. R., and Toral, S. 2022. “Prediction and Modelling Online Reviews Helpfulness Using 1D Convolutional Neural Networks,” Expert Systems with Applications (198), 116787.

Park, D.-H., and Lee, J. 2008. “eWOM Overload and Its Effect on Consumer Behavioral Intention Depending on Consumer Involvement,” Electronic Commerce Research and Applications (7:4), pp. 386-398.

KITS·ICEC·KIISS
2026 통합 국제학술대회

B1 Smart Logistics & Supply Chain Resilience

Session Chair : Fen Qin (Taiyuan Univ. of Technology)



B1

Sustainable Smart Logistics: A Systematic Review of Collaborative Mechanisms Enabled by Digital Technologies and the TOE Framework

Dongqing Huo (Anhui University of Science and Technology),Jinxin Yang (Hefei Technology College),
Chao Lu (Anhui University of Science and Technology)

To outline smart logistics research frontiers and development trajectories driven by digital technologies and their sustainability impacts, we synthesized 806 Web of Science articles (2011~2025) using bibliometric and systematic analyses. Research demonstrates the field's evolution from isolated applications to a cohesive Technology–Organization–Environment framework.We propose a sustainability–oriented collaborative model showing how technological empowerment, organizational transformation, and environmental constraints jointly foster sustainability. Specifically, IoT, AI, and decentralized networks□embedded in adaptive organizational processes and guided by policy–enhance resource efficiency, reduce emissions, and improve social welfare. This study positions smart logistics as a systemic sustainability enabler and offers practical insights for aligning digital transformation with sustainable value creation.

KEY WORDS

Smart Logistics, Sustainability, Sustainable Collaborative Mechanisms, Digital Technology Enablement

Sustainable Smart Logistics: A Systematic Review of Collaborative Mechanisms Enabled by Digital Technologies and the TOE Framework

Dongqing Huo [a], Jinxin Yang [b], and Chao Lu [c]
[a] School of Economics and Management, Anhui University of Science and Technology,
E-mail: huooodq@aust.edu.cn
[b] School of Economics and Trade, Hefei Technology College,
Email: jinxiny@aust.edu.cn
[c] The First Affiliated Hospital of Anhui University of Science and Technology,
E-mail: chaolu@aust.edu.cn

Abstract

Driven by digital transformation, smart logistics is evolving toward systemic sustainable collaboration. Following PRISMA guidelines, we systematically analyzed 806 Web of Science articles (2011–2025), tracing the shift from isolated technology adoption to integrated governance. We propose a TOE-based collaborative model demonstrating how digital empowerment, organizational restructuring, and environmental constraints jointly drive carbon reduction and operational efficiency. This study explicitly contributes a unified analytical framework, a transparent screening methodology, and validated implementation pathways. Findings underscore that sustainable smart logistics requires synchronized technological, organizational, and policy alignment.

Keywords: Smart Logistics, Sustainability, Digital Technology Enablement

Introduction

With the digital economy and "dual carbon" goals advancing, the logistics industry is undergoing digital and green transformation(Feng et al., 2025). Accounting for nearly 60% of global transport energy use, logistics faces urgent efficiency and sustainability challenges. Smart logistics employs IoT, AI, and blockchain to tackle high energy consumption and inefficiency(Tariq et al., 2023). Globally, policy and technology drive its development: Industry 4.0(Hofmann & Ruesch, 2017)introduced automated cyber-physical systems(Parhi et al., 2022), while Industry 5.0 (Golovianko et al., 2023)emphasizes human-machine collaboration and green innovation(Jefroy et al., 2022; Vudugula, 2025). EU policies like CBAM require carbon accounting, promoting digital decarbonization(Magacho et al., 2024). AI improves efficiency but raises governance issues(Guandalini, 2022), underscoring the need for a TOE framework(Wang et al., 2024).

Research indicates digital technologies enhance efficiency(Esmaeilian et al., 2020), but systematic studies on TOE collaboration are scarce(Woschank & Zsifkovits, 2021). Technological integration and interoperability remain inadequate, limiting synergies(Vudugula, 2025; Shee et al., 2021). Empirical studies on smart logistics' impact during digital transformation are lacking(Sang & Anh, 2025), as are comparative analyses of carbon policies' influence on technology adoption. This study

examines smart logistics trends and sustainability challenges via bibliometric analysis of 806 Web of Science articles (2011–2025). Using CiteSpace, it identifies a shift from isolated tech applications to TOE integration and proposes a model for digitally enabled sustainable smart logistics.

Co-citation Analysis Based on Cite Space

Research Methodology and Data Sources

This study uses bibliometrics to analyze digital technology in smart logistics, examining publication trends, author composition, and keyword frequency(Chen, 2017). Data from Web of Science with keywords "smart logistics," "digital technology," and "sustainability" from 2011-2025 yielded 806 articles. It integrates the TOE framework with CiteSpace for co-occurrence, clustering, timeline visualization, and burst detection to track thematic evolution and trends.

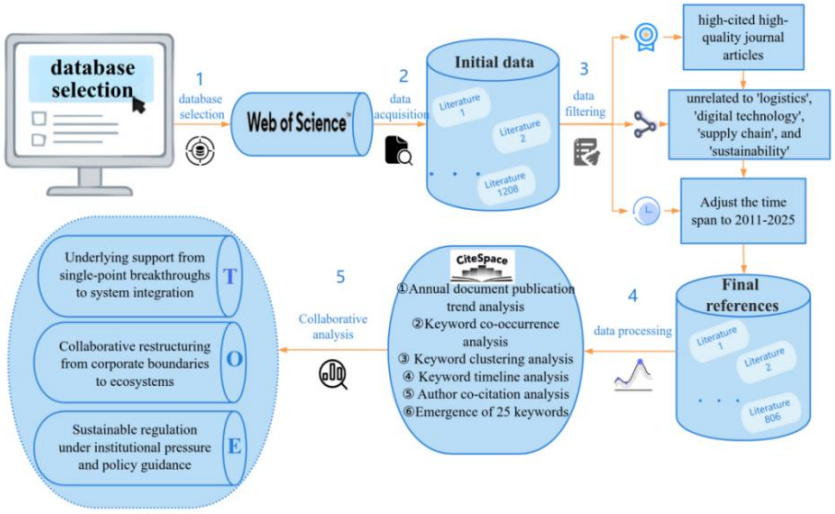


Figure 1. Research Model.

Annual Number of Publications Analysis

The number of published papers reflects latest research findings and academic interest, (Zhou et al., 2022)while its trend indicates research progress and knowledge (Zhang et al., 2025)maturity. Analyzing WOS data, we observe annual publication changes, gaining insights into activity in digital technology-enabled smart logistics research, as shown in Figure 2.

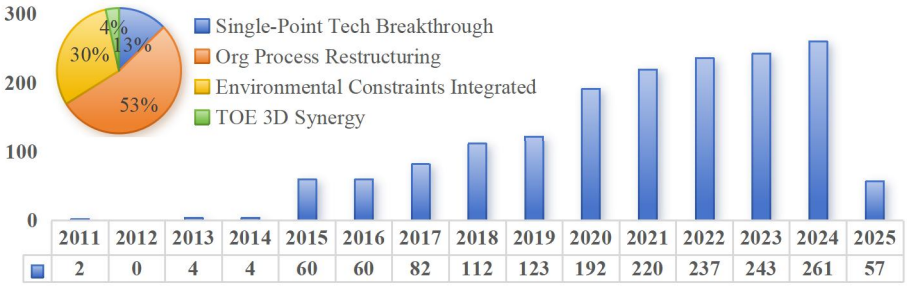


Figure 2. Publication Time Distribution of Academic Literature.

Figure 2 shows a sharp increase in publications from 2 in 2011 to 60 by 2015, driven by RFID and cloud computing. Growth reached 192 by 2020 with big data analytics and AI boosting academic interest, reflecting smart logistics' practical expansion. Post-2020, COVID-19 accelerated unmanned delivery, while regulations spurred carbon footprint research(Zhao et al., 2025). By 2025, steady growth indicates a shift toward sustainable implementation and cross-border collaboration. Thus,

smart logistics, fueled by digital technologies from 2011 to 2025, divides into four stages in Figure 3, analyzing drivers, challenges, transitions, and case studies.

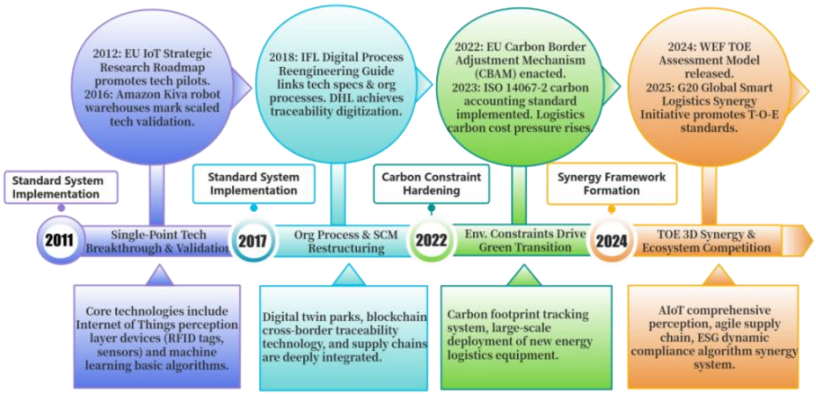


Figure 3. Research Stages of Smart Logistics Empowered by Digital Technology.

Keyword Co-occurrence Analysis

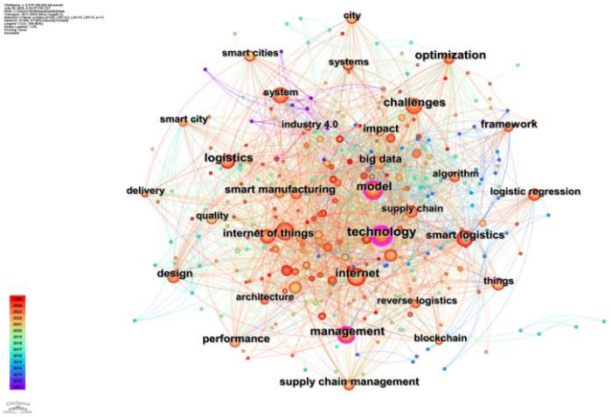


Figure 4. Keyword Co-occurrence Map.

Node size and link thickness represent term frequency and co-occurrence intensity, illustrating thematic evolution and interconnections(Santibanez Gonzalez et al., 2023). Node color gradients indicate temporal variations.

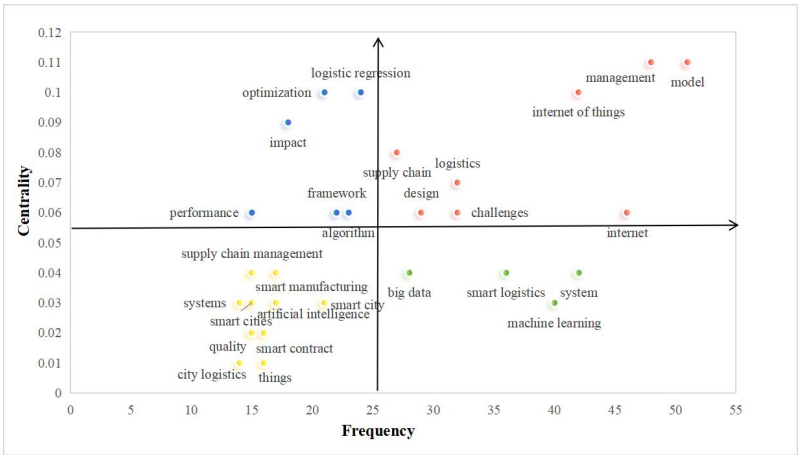


Figure 5. Strategic Coordinate Map of Keywords.

Technologically, IoT, ML, and AI are core, with IoT enabling logistics transparency via data collection, (Wang et al., 2025)and ML/AI (Ding et al., 2023)optimizing (N. et al., 2019)decisions like VRP.(Y. et al., 2020) Digital Twin and CPS show a move toward a cohesive tech ecosystem.

Organizationally, technology adoption depends on process reconfiguration, where Smart Contracts and decentralized networks reduce transaction costs, and VRP-ML links aid transportation scheduling. Smart manufacturing integrates with supply chain research, indicating greater integration. Environmentally, sustainability links to carbon traceability, embedding eco-goals into digital tech. The strategic diagram shows research shifting from isolated apps to comprehensive TOE-based studies(Callon et al., 1991).

Keyword Co-occurrence Timeline Map

Keyword clustering timeline (Figure 6) visualizes sustainable smart logistics evolution with digital technologies: node size indicates citation potential, dense branches show higher citation rates, line thickness reflects mutual citation strength, and color/weight denote publication year and frequency(Xu et al., 2023).Findings reveal strong links between digital transformation, R&D innovation, and smart logistics.Since 2013, darker nodes and thicker lines signify significantly accelerated research growth. Key emerging trends include the digital economy and platform-based(Centobelli et al., 2022) operations.

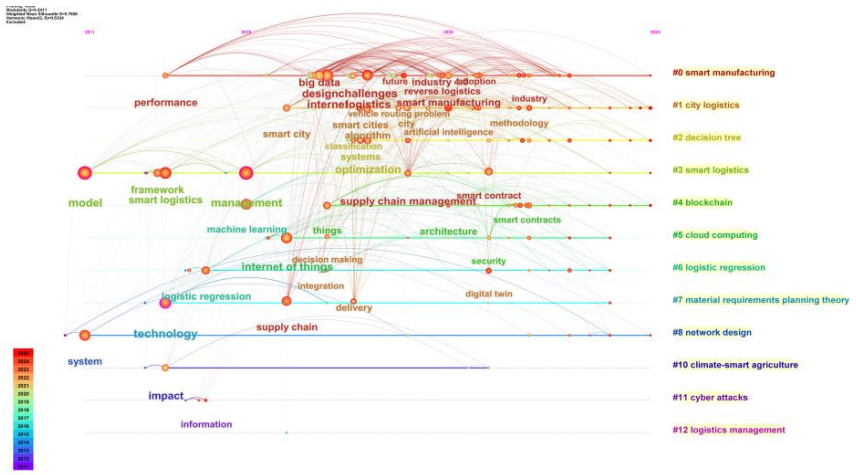


Figure 6. Co-linearity Time Map of Key Words.

Collaborative Mechanism Analysis Based on the TOE Framework

Smart logistics features three core themes: IT tools(Antonio et al., 2023; Winkelhaus & Grosse, 2020)(IoT, blockchain, big data, AI) such as IoT, blockchain, big data, and AI enhance visibility and decision-making; platform governance(Centobelli et al., 2022) regulates interactions and builds trust; policy adaptation(Jiabao et al., 2023) emphasizes regulations and standards(Chen et al., 2019). The TOE framework provides a systematic analytical approach due to its complexity: the technological dimension includes algorithms and data tools driving digital transformation(Centobelli et al., 2022); the organizational dimension involves process restructuring and platform collaboration; the environmental dimension covers policies and market dynamics enabling green transformation. Empirical studies confirm interactions: technology enables organizational change, environment drives adaptation, and multi-dimensional effects support sustainability. The goal is to achieve resource efficiency, sustainability, and social welfare(Xiu et al., 2025). This paper adds a value dimension to TOE, forming a "Technology-Organization-Environment-Value" model. This extension addresses the framework's explanatory limitations and shifts analysis from structural conditions to outcome-oriented perspectives, illustrating smart logistics' transition from functional optimization to sustainable collaboration.

Table 1. Comparative empirical study of intelligent logistics coordination mechanisms under the TOE framework

Author (Year)	Region	Data Source/Case	Key Findings and Coordination Mechanism	Dimension	Sustainable Benefits
(Abdullahi et al., 2025)	UK	Russell Transportation Group	Data intelligence (AI + federated learning) optimizes resource allocation; policy incentives for green transport	T+O+E	Fleet carbon emissions reduced by 20%-30%
(MoHURD, 2018)	Singapore	CSUS joint Laboratory Technical White Paper	"Digital carbon ticket" mechanism (Internet of Things + decentralized network) enables dynamic carbon emission monitoring, with government carbon credit subsidies motivating green process adaptation	T+E→O	Green transportation modal share has increased to 75%
(Lu et al., 2025)	China	Product carbon footprint scenario	Decentralized Network enhances supply chain carbon footprint accounting accuracy, responding to consumer green demands and policy compliance pressures	T→E	Carbon data accuracy of 92%
(Sunghun & Hyunsu, 2017)	South Korea	Amazon Prime Air service	Drones and ground robots collaborate to optimize last-mile delivery paths, reducing manual scheduling requirements	T→O	Net transportation cost of service is only 8.6% of that without the service
(Sharma et al., 2021)	India	Automotive parts industry cluster	Carbon emission policies force small suppliers to join cloud supply chain platforms, enabling organizational clustering	E→O	Supply chain carbon intensity decreased by 18%
(Ziatdinov et al., 2024)	Japan	Tokyo-Osaka freight line	5G vehicle network technology enables truck platoon collaborative driving, human-machine collaboration reduces driver fatigue	T+O→E	Highway accident rate decreased by 25%
(Yu et al., 2024)	China	A-share listed transport firms	Digital transformation boosts ESG (stronger in SOEs/large firms); financing constraints weaken this effect; green innovation mediates the link	T+O→E	Promotes low-carbon transformation, enhances social responsibility, eases financing constraints
(Yamchi et al., 2024)	India	Punjab agricultural supply chain	Farmer cooperatives collectively purchase solar-powered refrigerated trucks, responding to India's rural electrification policy	O→T+E	Fruit and vegetable loss rate reduced from 30% to 8%

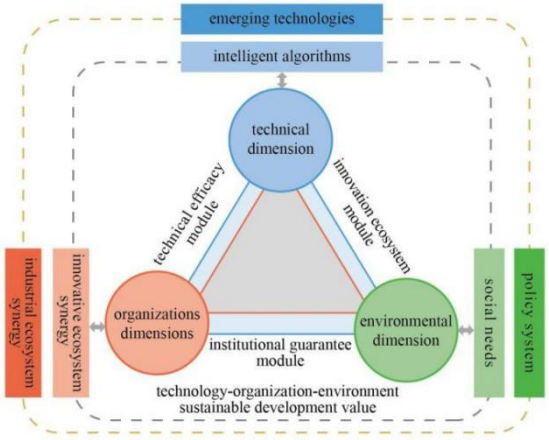


Figure 7. Smart Logistics Collaborative Mechanism Model

Discussion and Implications

Theoretical contributions introduce the TOE framework to address the lack of a system perspective in smart logistics, enhancing adoption theory with a "Technology-Organization-Environment-Value" structure that emphasizes a shift to systemic and value-oriented approaches focused on sustainable value. Practically, smart logistics requires technological, organizational, and institutional adaptability, but rapid evolution reveals research imbalances, such as models lacking dynamic paths and feedback loops, heterogeneity complicating policy, and platform data limiting empirical studies. Platform solutions like "labor intensity balance" and "dynamic pricing" integrate digital tools, while institutional tools like "digital carbon tickets" improve green governance. Future research should expand theories with Complex Adaptive Systems and Organizational Resilience, innovate methodologies, enhance data accessibility, balance efficiency, collaboration, and resilience, and address digital inequality and ethics through inclusive governance. In conclusion, smart logistics is evolving from localized optimization to system modeling and value orientation, with TOE-based collaboration key for sustainable advancement, requiring integrated efforts in theory, methodology, and practice.

References

Abdullahi, I., Larijani, H., Liarokapis, D., Paterson, J., Jones, D., & Murray, S. (2025). A Data-Intelligence-Driven Digital Twin Framework for Improving Sustainability in Logistics Applied Sciences (15, pp.).

Antonio, C., Elena, M. G., Vesna, V., & Carmen, C. (2023). Cross-industry innovation: A systematic literature review. Technovation, 124, 102743.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. Scientometrics, 22(1), 155-205.

Centobelli, P., Cerchione, R., Del Vecchio, P., Oropallo, E., & Secundo, G. (2022). Blockchain technology for bridging trust, traceability and transparency in circular supply chain. Information & Management, 59(7), 103508

Chen, C. (2017). Science Mapping:A Systematic Review of the Literature. Journal of Data and Information Science, 2(2), 1-40.

Chen, Y., Yin, Y., Browne, G. J., & Li, D. (2019). Adoption of building information modeling in Chinese construction industry The technology-organization-environment framework. Engineering Construction and Architectural Management, 26(9), 1878-1898.

Ding, G., Cui, L., Han, C., Wang, X., & Zhang, F. (2023). Simulation evaluation and analysis of aircraft group support based on multi-agent. Journal of Beijing University of Aeronautics and Astronautics, 49(9), 2306-2316.

Esmaeilian, B., Sarkis, J., Lewis, K., & Behdad, S. (2020). Blockchain for the future of sustainable supply chain management in Industry 4.0. Resources Conservation and Recycling, 163 ,105064

Feng, D., Wang, H., & Zhao, L. (2025). Digital Technologies for Sustainable Supply Chain Performance: Source-Push and Value Chain-Pull Mechanisms Sustainability (17, pp.).

Golovianko, M., Terziyan, V., Branytskyi, V., & Malyk, D. (2023). Industry 4.0 vs. Industry 5.0: Co-existence, Transition, or a Hybrid. Procedia Computer Science, 217, 102-113.

Guandalini, I. (2022). Sustainability through digital transformation: A systematic literature review for research guidance. Journal of Business Research, 148, 456-471.

Hofmann, E., & Ruesch, M. (2017). Industry 4.0 and the current status as well as future prospects on logistics. Computers in Industry, 89, 23-34. https://doi.org/10.1016/j.compind.2017.04.002

Jefroy, N., Azarian, M., & Yu, H. (2022). Moving from Industry 4.0 to Industry 5.0: What Are the Implications for Smart Logistics? Logistics (6, pp.).

Jiabao, L., Shunzhi, L., Jose, B., Xin, R. L., & Aseel, A. (2023). How to build supply chain resilience: The role of fit mechanisms between digitally-driven business capability and supply chain governance. Information &Amp; Management, 60(2), 103747.

Lu, Q., Liu, Y., Liu, C., Zeng, C., Chen, D., & Ning, J. (2025). Product Ownership Transfer Scheme of RFID-Enabled Supply Chain Based on Blockchain. Acta Electronica Sinica, 53(2), 451-459.

Magacho, G., Espagne, E., & Godin, A. (2024). Impacts of the CBAM on EU trade partners: consequences for developing countries. Climate Policy, 24(2), 243-259.

MoHURD. (2018). Centre for livable cities, Singapore smart city joint labs of CSUS. China Architecture & Building Press.

N., O. A., A., F. K., & A., M. (2019). Intelligent Transportation and Control Systems Using Data Mining and Machine Learning Techniques: A Comprehensive Study. Ieee Access, 7, 49830-49857.

Parhi, S., Joshi, K., Gunasekaran, A., & Sethuraman, K. (2022). Reflecting on an empirical study of the digitalization initiatives for sustainability on logistics: The concept of sustainable logistics 4.0. Cleaner Logistics and Supply Chain, 4, 100058.

Sang, T. M., & Anh, N. Q. (2025). Adopting digital technology for sustainable performance: circular economy practice in the transportation sector. Discover Sustainability, 6(1), 166.

Santibanez Gonzalez, E. D. R., Kandpal, V., Machado, M., Martens, M. L., & Majumdar, S. (2023). A Bibliometric Analysis of Circular Economies through Sustainable Smart Cities Sustainability (15, pp.).

Sharma, M., Kamble, S., Mani, V., Sehrawat, R., Belhadi, A., & Sharma, V. (2021). Industry 4.0 adoption for sustainability in multi-tier manufacturing supply chain in emerging economies. Journal of Cleaner Production, 281, 125013.

Shee, H. K., Miah, S. J., & De Vass, T. (2021). Impact of smart logistics on smart city sustainable performance: an empirical investigation. The International Journal of Logistics Managementthe International Journal of Logistics Management, 32(3), 821-845.

Sunghun, J., & Hyunsu, K. (2017). Analysis of Amazon Prime Air UAV Delivery Service. Journal of Knowledge Information Technology and Systems, 12(2), 253-266.

Tariq, S., Nasir, N., & Ali, A. (2023). Smart Logistics for Sustainable Supply Chain: A Bibliometric Analysis. Journal of Logistics Informatics and Service Science, 11, 284-306.

Vudugula, S. C. O. B. (2025). Sustainable smart supply chains: A review of green technologies and their impact on logistics. American Journal of Scholarly Research and Innovation, 4(01), 1-32.

Wang, S., Zhang, X., Peng, J., Tan, Y., & Fan, Z. (2024). Providing solutions for carbon emission reduction using the TOE framework. Expert Systems with Applications, 255, 124547.

Wang, X., Fei, Z., Zhou, Y., & Hu, J. (2025). Integrated Sensing, Communication, Computation, and Intelligence Towards IoT: Key Technologies and Future Directions. Journal of Electronics & Information Technology, 47(4), 888-908.

Winkelhaus, S., & Grosse, E. H. (2020). Logistics 4.0: a systematic review towards a new logistics system. International Journal of Production Research, 58(1), 18-43.

Woschank, M., & Zsifkovits, H. (2021). Smart Logistics-Conceptualization and Empirical Evidence. Chiang Mai University Journal of Natural Sciences, 20(2), 123-145.

Xiu, P., Yang, P., & Lin, X. (2025). Digital transformation and environmental, social and governance performance: a technology–organization–environment framework. Management Decision, 1-31.

Xu, C., Zhu, S., Yang, B., Miao, B., & Duan, Y. (2023). A Review of Policy Framework Research on Promoting Sustainable Transformation of Digital Innovation Sustainability (15, pp.).

Y., Z., Z., W., H., W., & F., B. (2020). Artificial Intelligence-Aided Thermal Model Considering Cross-Coupling Effects. Ieee Transactions On Power Electronics, 35(10), 9998-10002.

Yamchi, H. R., Jabarzadeh, Y., Govindan, K., & Mahdiraji, H. A. (2024). A triple bottom line approach for designing a sustainable closed-loop supply chain network in fruit industry: A metaheuristic solution approach. Journal of the Operational Research Society, 75(10), 1925-1948.

Yu, L., Xu, J., & Yuan, X. (2024). Sustainable Digital Shifts in Chinese Transport and Logistics: Exploring Green Innovations and Their ESG Implications Sustainability (16, pp.).

Zhang, L., Cui, Q., Zhang, L., & Hao, Y. (2025). Application trends of CGE models in low-carbon transition research: a bibliometric analysis. Journal of Beijing Normal University. Natural Science, 61(2), 178-190.

Zhao, H., Ye, Q., Zhang, Z., Zhang, K., Wang, K., & Huang, Z. (2025). Development of low carbon algorithms and application of compression and acceleration techniques. Control and Decision, 40(5), 1409-1428.

Zhou, J., Guo, Y., Sun, J., Yu, E., & Wang, R. (2022). Review of bike-sharing system studies using bibliometrics method. Journal of Traffic and Transportation Engineering (English Edition), 9(4), 608-630.

Ziatdinov, R., Atteraya, M. S., & Nabiyev, R. (2024). The Fifth Industrial Revolution as a Transformative Step towards Society 5.0 Societies (14, pp.).

B1

How Does Information Technology Innovation Promote the Sustainable Development of Logistics Enterprises? - A Systematic Review Based on the TOE Framework

Hailin Cao (Anhui University of Science and Technology), Jinxin Yang (Hefei Technology College),
Chao Lu (The First Affiliated Hospital of Anhui University of Science and Technology)

Based on the Technology–Organization–Environment (TOE) framework, this study reviews literature (2016~2025) to analyze how information technology (IT) innovation drives sustainable smart logistics. Moving beyond cost-centric views, it integrates environmental, economic, and social dimensions. Findings indicate that technology aids carbon reduction through visibility and optimization mechanisms. Organizational capabilities act as a bridge transforming tech into performance, while regulatory and market environments set boundary conditions. Crucially, the study identifies "data interoperability and trustworthy disclosure" as fundamental enablers. By redefining sustainable performance and explaining divergent implementation outcomes, this research broadens the TOE framework and offers actionable insights for integrating environmental goals into digital logistics under the dual-carbon strategy.

KEY WORDS

Smart Logistics, Information Technology Innovation, TOE Framework, Data Interoperability

How Does Information Technology Innovation Promote the Sustainable Development of Logistics Enterprises? — A Systematic Review Based on the TOE Framework

Hailin Cao [a], Jinxin Yang [b], and Chao Lu [c]
[a] School of Economics and Management, Anhui University of Science and Technology,
E-mail: caohailin@aust.edu.cn
[b] School of Economics and Trade, Hefei Technology College,
Email: jinxiny@aust.edu.cn
[c] The First Affiliated Hospital of Anhui University of Science and Technology,
E-mail: chaolu@aust.edu.cn

Abstract

Based on the Technology–Organization–Environment (TOE) framework, this study reviews literature (2016–2025) to analyze how information technology (IT) innovation drives sustainable smart logistics. Moving beyond cost-centric views, it integrates environmental, economic, and social dimensions. Findings indicate that technology aids carbon reduction through visibility and optimization mechanisms. Organizational capabilities act as a bridge transforming tech into performance, while regulatory and market environments set boundary conditions. Crucially, the study identifies "data interoperability and trustworthy disclosure" as fundamental enablers. By redefining sustainable performance and explaining divergent implementation outcomes, this research broadens the TOE framework and offers actionable insights for integrating environmental goals into digital logistics under the dual-carbon strategy.

Keywords: Smart Logistics; Information Technology Innovation; TOE Framework; Data Interoperability

Introduction

Driven by global supply chain instability and carbon neutrality goals, logistics has evolved from a cost-driven support activity into a core platform for value creation, resilience, and long-term competitiveness. IT advancements, such as the Internet of Things (IoT), artificial intelligence (AI), blockchain, and digital twins provide unprecedented transparency and control over logistics operations. This reduces empty runs and energy use while improving load efficiency, safety compliance, and service reach. However, technological progress alone cannot guarantee sustainable outcomes. Technologies must be deeply integrated with an organization's internal capabilities and external environment to embed "greenness" and "resilience" into everyday operations rather than treating them as short-term projects.

Existing literature on IT innovation and performance often focuses narrowly on technological efficiency and instrumental logic. It frequently neglects the mediating roles of organizational absorptive capacity, process redesign, and cross-departmental data management. Furthermore, studies on external pressures like government policy or market competition rarely explain how these forces translate into concrete operational practices. Consequently, two paradoxes persist in practice: firms adopting similar technologies achieve divergent performance outcomes, and successful green pilot projects often lose momentum when scaled. These issues frequently stem from a lack of data interoperability and reliable disclosure, preventing the stable integration of verifiable green initiatives into organizational routines, incentive systems, and cross-boundary coordination

To address these gaps and move beyond the static categorization typical of previous TOE literature, this research re-conceptualizes the TOE framework as a dynamic and interactive system. We expand the dependent variable from pure economic performance to encompass environmental, economic, and social sustainability. Unlike traditional studies that treat T, O, and E as isolated adoption factors, we propose a synergistic mechanism. First, Technology builds a traceable and visible data loop through sensing, optimization, and substitution. Second, Organization acts as a capability translation engine. Governance alignment and ambidextrous capabilities translate raw technological data into systemic process improvements. Third, Environment serves as an external catalyst. Regulatory and market pressures provide incentives and boundary conditions for wider diffusion. Together, these dynamic mechanisms shift green objectives from isolated visions to routine practices. The study addresses three questions. (1) how IT innovation impacts sustainable performance. (2) how organizational governance mediates this relationship. (3) how external environments moderate technology-organization interactions to enable supply chain diffusion.

Methodologically, this research combines a systematic literature review with quantitative Literature Analysis using Web of Science Core Collection data (2016 – 2025). By analyzing keyword co-occurrence, clustering, and thematic development, this approach balances comprehensive coverage with deep interpretation. It translates abstract network patterns into practical insights, ensuring discussions remain evidence-based while offering scenario references for decision-making.

This study contributes on three levels. Theoretically, it extends the TOE framework to a multi-dimensional sustainability system and identifies data interoperability and trustworthy disclosure as key mechanisms to explain the "similar technologies, divergent outcomes" and "pilot success–scaling decay" phenomena. Methodologically, it integrates literature mapping with content analysis to construct a verifiable " data – capability – performance " chain. Practically, it provides actionable pathways to embed green metrics—such as carbon intensity, energy per order, and reverse recovery rates—into daily warehouse management system (WMS), transportation management system (TMS), and order management system (OMS) reporting systems, leveraging platform governance to diffuse sustainable capabilities across supply chain networks.

Research Methods

Systematic Literature Review Integrating Quantitative Literature Analysis

To investigate how IT innovation improves business performance for logistics service providers (LSPs), this study integrates a systematic literature review with quantitative literature analysis. The structured, three-step procedure involves: (1) gathering bibliographic data, (2) creating visualization maps using CiteSpace, and (3) interpreting the intellectual structure and core topics through the lens of the TOE framework. To move beyond macro-level patterns and understand the underlying mechanisms, this quantitative approach is complemented by an in-depth content analysis of 31 highly cited, relevant articles.

Data Sources and Refinement

Drawing on the methodology of Zupic et al, this study conducted a systematic search in the Web of Science Core Collection. The search query “Logistics Service Provider” AND (“IT innovation” OR “information technology innovation”) AND “Business performance” was applied to titles, abstracts, and keywords, initially yielding 2,107 documents. Because IT innovation in logistics is highly multidisciplinary, the scope was refined to focus strictly on management and operations. Accordingly, the results were filtered to include only five subject categories: Management, Business, Industrial Engineering, Operations Research & Management Science, and Transportation. By further restricting the document types to articles and review papers (excluding conference papers) published between 2016 and 2025, the dataset was finalized at 928 publications. The complete bibliographic records of these documents were then exported for subsequent analysis.

Data Analysis

Core Theme Identification Based on High-Centrality Keywords

Centrality indicates how keywords act as "bridges" within the co-occurrence network. In the network illustrated in Figure 1 the median centrality is around 0.02, so keywords with a centrality of 0.05 or higher can be considered the main themes of the field. As presented in Table 1, these keywords play a particularly important role in the research, highlighting the domain's core topics.

Nodes with centrality of 0.05 or higher function as bridges across different themes. Among these, "information technology capability" (0.08, 2016) has the highest centrality, showing that the IT capability framework serves as a central hub connecting technological investment and performance. "Business performance" (0.05, 2016) reflects the main research focus on corporate outcomes. The keywords "adoption" (0.04, 2016), "integration" (0.04, 2016), and "green innovation" (0.05, 2019) highlight the causal links across process, context, and sustainability dimensions. Together, these high-centrality nodes outline the dominant logic of "IT–Capability–Performance."

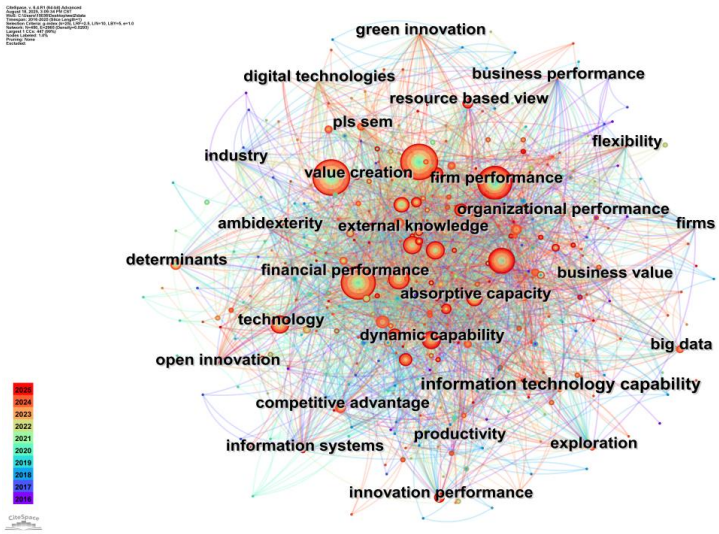


Figure 1. Keyword Co-occurrence Map

Table 1. High-Centrality Keywords

Keyword	Centrality	Year	Keyword Indication
IT capability	0.08	2016	IT capability is the core mediating variable linking IT innovation and performance
Business performance	0.05	2016	Reflects the research objective and forms an input–output chain with IT innovation.
adoption	0.04	2016	Technology adoption initiates IT innovation's impact on performance.
integration	0.04	2016	IT integration with processes and supply chains is key to value realization.
green innovation	0.05	2019	Emerging theme showing sustainable technologies' impact on performance.

Keyword clustering analysis under the TOE framework.

This study employs keyword clustering to map the intellectual structure of the field, a method that reduces the complexity of information while offering a systematic lens through which researchers can comprehend and contextualize its knowledge base. We applied this approach to a decade of authoritative international literature, using Cite Space to generate a co-word network (Figure 2). The

resultant clusters are then interpreted within the established Technology-Organization-Environment (TOE) framework to delineate core research trajectories.

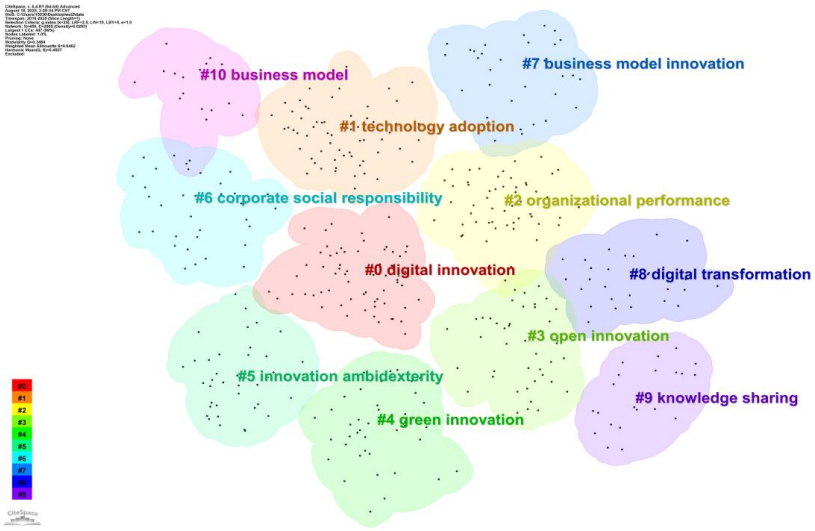


Figure 2. word Cluster Map

By deconstructing the keyword clustering map (Figure 2), this study categorizes nine keyword clusters into the three dimensions of the TOE framework—Technology (T), Organization (O), and Environment (E)—as shown in Table 4, and further reveals their underlying logical interconnections.

Keyword analysis under the three dimensions of the TOE framework

Table 2. Keyword Clusters under the TOE Framework

Dimension	Cluster Number & Keyword	Core Connotation
Technology (T)	#0 digital innovation	Application of IoT, AI, and other new technologies in logistics
	#1 technology adoption	Decision process of introducing IT in logistics firms
	#8 digital transformation	Technology-driven restructuring of logistics systems
Organization (O)	#5 innovation ambidexterity	Balancing efficiency and disruptive innovation
	#3 open innovation	Integrating external resources for joint R&D
	#9 knowledge sharing	Transferring IT experience across departments or firms
	#10 business model	Transformation of value creation and capture by IT innovation
Environment (E)	#4 green innovation	Technological change driven by policies and sustainability demands
	#6 corporate social responsibility	Social expectations guiding technological innovation

Based on keyword clustering (Table 2, Figure 2), the Technological dimension (T) acts as the primary driver of performance. Clusters #0 “digital innovation” and #8 “digital transformation” represent core applications (IoT, AI) and strategic changes. Notably, Cluster #1 “technology adoption”) holds the

highest log-likelihood ratio (LLR) value (23.46), emphasizing that performance gains depend less on the technology itself and more on successful implementation processes.

The Organizational dimension (O) serves as the critical mediator. Cluster #5 " ambidextrous innovation " strategically bridges technological tools #0 with environmental goals #4 " green innovation". This mediating role is reinforced by "knowledge sharing" #9 and "open innovation" #3, which facilitate resource integration. Ultimately, organizational capabilities determine how effectively technological inputs are translated into performance outputs.

The Environmental dimension (E) establishes boundary conditions and catalysts. Cluster #4 "green innovation" and #6 "corporate social responsibility" impose policy pressures and normative standards, steering the direction of digital innovation toward sustainability. In synthesis, Technology supplies the instrumental pathways, the Organization acts as the adaptive engine, and the Environment sets the driving rules, confirming the dynamic relevance of the TOE framework.

Journal Co-Citation Analysis

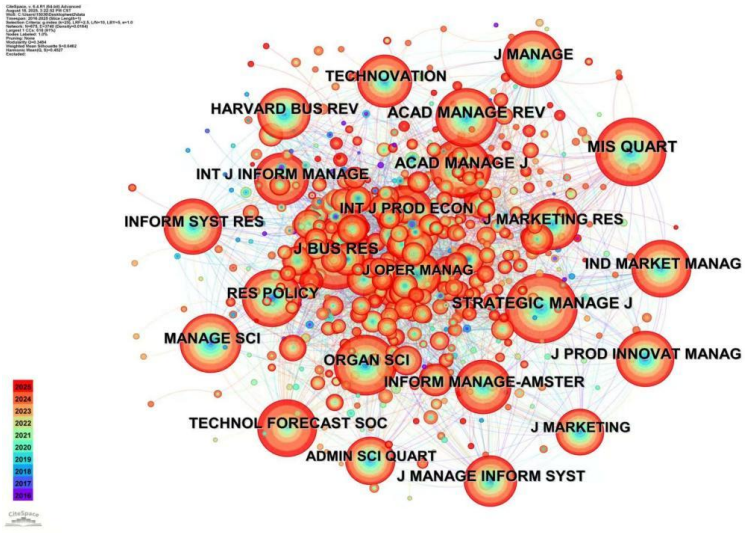


Figure 3. Journal Co-citation Visualization

Analyzing the journal co-citation network allows for a clear identification of the main knowledge sources and the interdisciplinary context that support this research field. The results show that the field's knowledge base is diverse and integrates multiple disciplines, with four main clusters: strategic management, information systems and technology innovation, marketing, and operations and supply chain management. Based on the selected literature, this study employs Cite Space software to generate the journal co-citation visualization, as shown in Figure 3, and Table 3 clearly presents the top 15 journals with the highest co-citation frequencies.

Table 3. Top 15 Most Frequently Cited Journals

No.	Journal Name	Frequency	Journal Type
1	STRATEGIC MANAGE J	640	Management
2	MIS QUART	623	Information Systems
3	ACAD MANAGE REV	482	Management
4	ACAD MANAGE J	480	Management
5	J MANAGE	468	Management
6	ORGAN SCI	462	Organizational Science
7	TECHNOL FORECAST SOC	459	Technology Management
8	MANAGE SCI	449	Management

9	IND MARKET MANAG	443	Marketing
10	INFORM SYST RES	431	Information Systems
11	INT J PROD ECON	421	Operations Management
12	INFORM MANAGE AMSTER	405	Information Management
13	RES POLICY	404	Policy Studies
14	HARVARD BUS REV	376	Business Review
15	J PROD INNOVAT MANAG	366	Innovation Management

The research landscape is anchored by four key pillars. Strategic Management journals provide the theoretical foundation. Information Systems journals, led by MIS Quarterly (623 citations) and Information Systems Research, constitute a major pillar emphasizing IT as a critical enabler. Marketing and Business journals play a central role, with the Journal of Business Research (656 citations) ranking first, highlighting the importance of market orientation. Finally, Operations journals contribute insights on efficiency and optimization. Collectively, this high-centrality co-citation network underscores the field's interdisciplinary nature, where intersections across these domains drive innovation and signal emerging research trends.

Literature Analysis under the TOE Framework

Based on quantitative literature and content analysis, this study finds that IT innovation in smart logistics drives multidimensional sustainability (environmental, economic, and social) not in isolation, but through the synergistic effects of the TOE framework. The following sections detail each dimension, their interactive mechanisms, and boundary conditions.

Technological Dimension: Digitalization-Based Pathways to Green Transformation

Technological adoption follows a stepwise path of visibility, optimization, substitution, and closed-loop management. Technologies like IoT and AI enhance system transparency to allow real-time monitoring of carbon emissions. Advanced AI routing minimizes resource waste, while blockchain traceability embeds sustainability into daily operations.

Organizational Dimension: Accumulation and Diffusion of Green Capabilities

Technology merely provides potential, whereas organizational capabilities act as the engine for translation. To achieve lasting performance, firms must execute several key actions. They need to incorporate green metrics into corporate governance and performance evaluations. They must enhance absorptive capacity to embed new practices into routine operations. Furthermore, they should develop organizational ambidexterity to balance daily efficiency improvements with exploratory green pilots.

Environmental Dimension: Directed Pressure Shaped by Policy and Market

The external environment acts as both a "directed pressure" and "diffusion accelerator." Policies like carbon pricing and green transport certifications shift sustainable practices from voluntary initiatives to mandatory compliance. Simultaneously, customer and investor demands for environmental, social, and governance (ESG) disclosures force firms to implement supply chain-wide carbon accounting and traceability systems, accelerating the standardization of green data governance.

Three-Dimensional Interaction and Boundary Conditions: The Critical Role of Data Interoperability

The TOE dimensions interact to link data, capabilities, and performance: the environment sets green targets, technology creates a continuous data loop, and organizational governance translates data into optimized routines. Data interoperability is the fundamental enabler here, allowing green outcomes to be measured, rewarded, and scaled. However, this mechanism is subject to boundary conditions. A

firm's asset structure, business scenarios, and regional policies dictate how technological and organizational measures are balanced and implemented.

Discussion and Implications

Theoretical Advancement

This study systematically examines smart logistics using the TOE framework, centering on environmental, economic, and social sustainability rather than just cost or efficiency. It emphasizes the interaction among the three dimensions. Technologically, data visibility and traceability enhance both efficiency and performance measurement. Organizationally, knowledge absorption and ambidextrous capabilities embed green objectives into routine operations, transforming potential into consistent practice. Environmentally, policy and market pressures shift green initiatives from optional to mandatory compliance. Across all dimensions, standardized data and reliable disclosure serve as pivotal mechanisms, enabling accurate measurement and supply chain diffusion. This explains phenomena like "divergent outcomes from identical technologies." Consequently, the study highlights data interoperability as both a mediator and threshold factor, guiding the TOE framework toward dynamic capabilities and verifiable, multi-dimensional performance.

Managerial Implications

From a managerial perspective, the core implication is to transform green goals from strategic aspirations into routine practices. Managers should integrate sustainability metrics alongside traditional cost and timeliness indicators. To operationalize this, firms must leverage technologies like IoT, AI, and blockchain for real-time performance assessment, supported by cross-departmental data sharing and staff training. Furthermore, successful pilot projects should be systematically scaled and embedded directly into IT systems and daily reporting. This ensures sustainability becomes a permanent, institutionalized component of business operations rather than a temporary initiative.

Policy Implications

The green transition of smart logistics requires robust policy and industry support alongside corporate efforts. Policymakers should design context-specific regulations, such as emphasizing energy efficiency for cold chains or low-emission zones for urban distribution. To lower cross-enterprise integration costs, regulators and industry associations must establish standardized carbon accounting and blockchain-based trusted data platforms. Furthermore, to ensure an equitable transition, governments and capital markets should provide financial aid, green loans, and supply chain financing to help SMEs upgrade their green infrastructure. Finally, cross-regional initiatives like “ green corridors ” can generate reproducible models. Ultimately, policy frameworks must act as both regulatory constraints and critical enablers for scaling sustainable capabilities.

Limitations and Future Research

While this study integrates sustainable smart logistics within the TOE framework, it relies primarily on literature and cases, limiting the assessment of underlying mechanisms. Furthermore, regional differences and the dynamic nature of green capability development are difficult to capture through cross-sectional designs. To address these limitations, future research should develop multi-dimensional databases combining operational and financial metrics, employ causal inference methods like quasi-experiments to validate data interoperability, and theoretically explore how asset structures and institutional environments influence the balance between low-carbon goals and resilience, ultimately moving the field from descriptive frameworks to rigorous empirical validation.

References

Alam, K., Ali, M.A., Erdiaw-Kwasie, M., Shahiduzzaman, M., Velayutham, E., Murray, P.A., and Wiesner, R. 2022. "Impact of ICTs on Innovation and Performance of Firms: Do Start-ups, Regional Proximity and Skills Matter?" *SUSTAINABILITY* (14:580110).

Bajaher, M.S., and Asiri, M. 2025. "The impact of digital transformation on ESG outcomes in the Saudi stock market," *Discover Sustainability* (6:1), p 752.

Abourokbah, S.H., Mashat, R.M., and Salam, M.A. 2023. "Role of Absorptive Capacity, Digital Capability, Agility, and Resilience in Supply Chain Innovation Performance," *SUSTAINABILITY* (15:36364).

Bogetoft, P., Kroman, L., Smilgins, A., and Sørensen, A. 2024. "Innovation strategies and firm performance," *Journal of Productivity Analysis* (62:2), pp 175-196.

Cao, G., Duan, Y., and Edwards, J.S. 2025. "Organizational culture, digital transformation, and product innovation," *INFORMATION & MANAGEMENT* (62:1041354).

Centobelli, P., Cerchione, R., and Esposito, E. 2018. "Environmental Sustainability and Energy-Efficient Supply Chain Management: A Review of Research Trends and Proposed Guidelines," in: *Energies*.

Chen, T., Cheng, H.K., Jin, Y.J., Li, S., and Qiu, L. 2021. "Impact of Competition on Innovations of IT Industry: An Empirical Investigation," *JOURNAL OF MANAGEMENT INFORMATION SYSTEMS* (38:3), pp 647-666.

Cui, T., Ye, J.H., and Tan, C.H. 2022. "Information technology in open innovation: A resource orchestration perspective," *INFORMATION & MANAGEMENT* (59:1036998).

Durugbo, C.M. 2025. "Uncertainties of supply chain transition and integration for Industry 5.0: a review of operations management literature," *International Journal of Production Research* (63:18), pp 6915-6937.

Ciasullo, M.V., Chiarini, A., and Palumbo, R. 2024. "Mastering the interplay of organizational resilience and sustainability: Insights from a hybrid literature review," *Business Strategy and the Environment* (33:2), pp 1418-1446.

Galkin, A., Samchuk, G., Kopytkov, D., and Thompson, R.G. 2025. "Digital twins in logistics: a comprehensive bibliometric analysis for advancing smart cities and sustainable development," *Discover Sustainability* (6:1), p 853.

Gao, Q., Xiao, J., Cao, Y., Deng, S., Ouyang, C., and Feng, Z. 2023. "Blockchain-based collaborative edge computing: efficiency, incentive and trust," *Journal of Cloud Computing* (12:1), p 72.

Han, N., and Um, J. 2024. "Risk management strategy for supply chain sustainability and resilience capability," *Risk Management* (26:2), p 6.

Castillo, A., Ruiz, L., Benitez, J., Schuberth, F., and Reina, R. 2023. "IT impact on open innovation performance: Insights from a large-scale empirical investigation," *DECISION SUPPORT SYSTEMS* (175:114025).

Chaudhuri, R., Chatterjee, S., Gupta, S., and Kamble, S. 2023. "Green supply chain technology and organization performance: Moderating role of environmental dynamism and product-service innovation capability," *TECHNOVATION* (128:102857).

Khuntia, J., Saldanha, T., Kathuria, A., and Tanniru, M.R. 2024. "Digital service flexibility: a conceptual framework and roadmap for digital business transformation," *EUROPEAN JOURNAL OF INFORMATION SYSTEMS* (33:1), pp 61-79.

Cuñat Negueroles, S., Reinosa Simón, R., Julián, M., Belsa, A., Lacalle, I., S-Julián, R., and Palau, C.E. 2024. "A Blockchain-based Digital Twin for IoT deployments in logistics and transportation," *Future Generation Computer Systems* (158), pp 73-88.

Liczmanska-Kopcewicz, K., Wisniewska, A., and Nocella, G. 2024. "Willingness to implement innovative solutions for creating information-based added value in food value chains," *JOURNAL OF CLEANER PRODUCTION* (446:141284).

B1

AI Applications in Data Analysis and Automated Decision-Making in Logistics Smart Contracts.

Thu Uyen Ha (University of Hull)

This study investigates the integration of Artificial Intelligence (AI) in logistics smart contracts, focusing on its effectiveness in enhancing data analysis and automating decision-making processes within blockchain-based logistics systems. By applying a mixed-methods approach, the research combines quantitative data from 120 real smart contract transactions on a decentralized logistics platform with qualitative insights from five industry experts. Experts emphasized that while AI-supported smart contracts offer clear operational benefits, successful deployment requires investments in digital infrastructure, workforce training, and clear regulatory frameworks. The study contributes to a growing body of knowledge on blockchain-AI convergence and provides practical policy implications for logistics firms, technology developers, and regulators aiming to scale AI enabled smart contract solutions. The study include the creation of AI governance standards, targeted incentives for innovation adoption, and the establishment of collaborative testbeds for safe experimentation and deployment.

KEY WORDS

Artificial Intelligence, Smart Contracts, Logistics, Data Analytics, Automated Decision-Making

ICEC 2026 Paper Title

AI Applications in Data Analysis and Automated Decision-Making in Logistics Smart Contracts

Thu Uyen Ha

Department of Marketing Management And Business Strategy, University of Hull

E-mail: Thuuyenha218@gmail.com

Abstract

This study investigates the integration of Artificial Intelligence (AI) in logistics smart contracts, focusing on its effectiveness in enhancing data analysis and automating decision-making processes within blockchain-based logistics systems. By applying a mixed-methods approach, the research combines quantitative data from 120 real smart contract transactions on a decentralized logistics platform with qualitative insights from five industry experts. Experts emphasized that while AI-supported smart contracts offer clear operational benefits, successful deployment requires investments in digital infrastructure, workforce training, and clear regulatory frameworks. The study contributes to a growing body of knowledge on blockchain-AI convergence and provides practical policy implications for logistics firms, technology developers, and regulators aiming to scale AI-enabled smart contract solutions. The study include the creation of AI governance standards, targeted incentives for innovation adoption, and the establishment of collaborative testbeds for safe experimentation and deployment.

Keywords: Artificial Intelligence; Smart Contracts; Logistics; Data Analytics;Automated Decision-Making

1. Introduction

The rapid digitalization of global logistics networks has necessitated the integration of intelligent technologies to address challenges such as information asymmetry, operational inefficiencies, and lack of transparency. Amid these transformations, blockchain-based smart contracts have emerged as pivotal instruments for enhancing trust and automation in logistics and supply chain management. However, the complexity of logistics operations, characterized by multidimensional data flows and dynamic decision environments, calls for advanced computational methods to support real-time analysis and autonomous decision making. In this context, artificial intelligence (AI) serves as a complementary technology that extends the functional capacities of smart contracts by enabling adaptive analytics, intelligent decision modelling, and predictive capabilities.

However, smart contracts alone lack the cognitive ability to derive insights from vast logistics datasets, which are critical for

decision-making. The integration of AI with smart contract mechanisms provides a promising solution to these limitations by embedding learning capabilities and data-driven reasoning into autonomous execution protocols (Virovets et al., 2024).

2. Literature review

The convergence of artificial intelligence (AI), blockchain technology, and smart contracts has transformed the landscape of logistics and supply chain management by facilitating data-driven decision-making, automating transactions, and enhancing transparency throughout supply chain processes. This literature review provides an in-depth examination of current academic and industrial research on AI applications in data analysis and automated decision-making within logistics smart contracts. Key themes include the integration of AI with blockchain, the design and execution of intelligent smart contracts, data analytics for real-time logistics decision-making, and privacy and security in decentralized systems. This paper lacks empirical information regarding socio-cultural disparities in automated contractual systems to elucidate artificial intelligence systems, hence influencing stakeholders' perspectives of AI-driven systems.

2.1 Integration of artificial intelligence and blockchain in logistics

Recent studies emphasize the increasing demand for decentralized and intelligent logistics systems capable of operating autonomously and securely. Abdelhamid, Sliman, and Ben Djemaa (2024) explored an AI-enhanced blockchain model tailored for scalable Internet of Things (IoT)- based supply chains. This research adopted a questionnaire survey involving 281 managers, with the aim of comprehensively examining the current state of AI integration across the U.S. supply chain sector, with a focus on some key components like real-time tracking, cost optimization, and risk management.

Virovets et al. (2024) further advanced this integration by introducing cryptographic oracles, which serve as trusted bridges between AI systems and blockchain-based smart contracts. These oracles facilitate secure, real-time data flows from AI modules to contract logic, ensuring that the smart contracts execute decisions based on verified analytics outcomes.

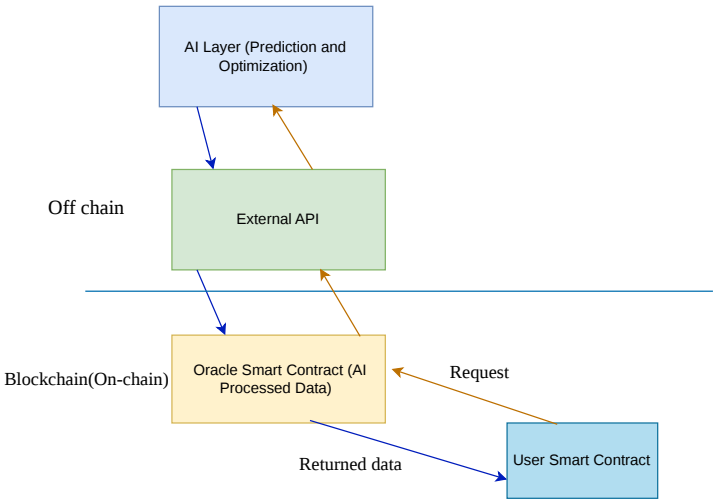


Figure 1: Traditional interaction of AI with smart contracts via an oracle

2.2 Smart contracts in supply chain logistics

Smart contracts are blockchain-based applications that, under certain circumstances, automatically carry out and enforce predetermined contractual obligations, which eliminates the need for middlemen and increases transactional transparency. Within the logistics domain, these contracts enable process automation, conditional payment settlements, and real-time compliance enforcement. Alqarni et al. (2023) examined the role of blockchain-based smart contracts in streamlining logistics operations. Their findings indicated that by embedding business logic directly into the contract, processes such as cargo tracking, proof of delivery, and customs clearance could be executed automatically.

2.3 Data analysis and AI-Driven decision making

The growing amount, variety, and urgency of logistics data exceed the capabilities of traditional analysis methods, necessitating more advanced and adaptable analytical abilities. Artificial intelligence (AI) techniques such as machine learning (ML), deep learning (DL), and natural language processing (NLP) have demonstrated their ability to identify patterns and draw valuable conclusions from vast volumes of diverse logistical data. Predictive analytics support route optimization, demand forecasting, and anomaly detection, which feed into smart contracts for automated adjustments in service-level agreements (SLAs) or delivery conditions..

2.4 Frameworks and models for logistics automation

The use of AI in smart contract-based logistics has been supported by a number of conceptual and practical models, especially in the areas of automated decision-making, risk assessment, and process optimisation. Groschopf, Dobrovnik, and Herneth (2021) introduced a maturity model for evaluating supply chain sustainability

using smart contracts. Their framework assesses whether logistics contracts are capable of capturing ESG (Environmental, Social, Governance) metrics and adapting operational decisions based on sustainability goals. autonomous material handling and delivery scheduling. This integration ensures real-time alignment between production activities and supply chain movements, minimizing downtime and inventory discrepancies.

2.5 Challenges and future directions

Despite their tremendous advancements, several organisational and technological obstacles still stand in the way of the efficient application of AI and smart contracts in logistics. In addition, interoperability between different blockchain platforms remains a hurdle. AI modules that support logistics decisions may need to interact with contracts deployed across Ethereum, Hyperledger Fabric, or other decentralized platforms. A unified protocol for AI-smart contract communication is essential to standardize decision-making across systems. This requirement calls for the incorporation of explainable AI (XAI) principles into the smart contractt ecosystem. Finally, the energy consumption and environmental impact of AI-blockchain integration cannot be overlooked. Through secure, privacy-preserving, and decentralized architectures, logistics networks are becoming more intelligent and adaptive. However, continued research is necessary to address scalability and ethical considerations, ensuring that these systems fulfil their potential in real-world logistics operations.

3. Data and research methods

This research employs a mixed-methods approach, integrating quantitative analysis of 20 smart contract transactions from a blockchain logistics platform with qualitative interviews with 5 industry professionals. Machine learning methods, including decision trees and neural networks, were employed to assess performance parameters such as execution time, cost, and mistake rate. Regression and clustering methodologies revealed substantial connections between AI integration and logistical efficiency. Thematic analysis of expert interviews revealed many challenges regarding adoption and significant strategic elements influencing the implementation of AI-enabled smart contract systems. The aim of using this strategy was to obtain a varied cohort exceeding 120 people from various demographic backgrounds. To guarantee data transparency, replies were evaluated for this strategy.

Data collection: A total of 120 transactions were collected from the blockchain system (https://etherscan.io) of a logistics platform utilising smart contracts. The data encompasses factors including execution time (in seconds), transaction cost (gas fee), error rate (number of unsuccessful transactions), and the degree of AI engagement in each transaction. Z-score normalisation was utilised for execution time and gas expenditure. Transactions that above three standard deviations from the mean were identified as anomalies and omitted from the regression analysis. This produced a pristine dataset of 112 transactions. AI interaction levels were represented as ordinal values (0 to 3), facilitating their incorporation into quantitative

*models.*Before analyzing the dataset, it was pre-processed to make sure the data was of good quality and consistent. This included dealing with missing values, normalizing key variables, and getting rid of outliers. Descriptive statistics were also examined in order to identify overarching tendencies in the data.

4. Findings and discussion

4.1 Regression analysis of AI interaction and logistics performance

A multiple linear regression model was developed to analyse the correlations between AI interaction levels and logistical performance variables. The results indicated a statistically significant negative association between AI interaction and execution time ($p < 0.01$), implying that more AI support decreased contract execution time. Transaction cost had a little negative correlation with AI level ($p = 0.041$), whereas mistake rate shown a significant negative correlation ($p < 0.001$).

Table 1. Regression coefficients: AI level vs. transaction metrics

Dependent Variable	Coefficient (AI Level)	Std Erro	t-Valu	p-Value
Execution Time	-3.215	0.842	-3.82	0.0002
Transaction Cost	-0.00037	0.00018	-2.05	0.041
Error Rate	-0.643	0.132	-4.87	<0.0001

These data indicate that the use of AI in smart contracts significantly enhances operational efficiency. Decreased execution time indicates enhanced throughput in logistics networks, but diminished mistake rates imply more accuracy in fulfilment and increased dependability of automated judgements. A k-means clustering approach was employed to categorise transactions into clusters based on execution time, cost, mistake rate, and AI level, in order to elucidate the impact of various AI adoption patterns on performance. The Elbow approach identified three as the best number of clusters.

4.2 Machine learning model performance in predicting transaction outcomes

Two machine learning models were trained to predict transaction outcomes based on input features including AI level, node latency, and transaction volume: a decision tree (CART) and a neural network (MLP). Evaluation of the models used an 80/20 train-test split. Root mean square error (RMSE) and R^2 values were calculated to assess predictive accuracy.

Table 2: Model performance metrics

Model	Predicted Variable	RMSE	R^2

Decision Tree	Execution Time	4.76	0.71
Neural Network	Execution Time	2.19	0.87
Decision Tree	Error Rate	0.66	0.68
Neural Network	Error Rate	0.34	0.91

Neural networks outperformed decision trees in all metrics. This indicates that non-linear models capture complex dependencies among variables in blockchain-based logistics environments. Such results support the implementation of deeper AI systems in future smart contract frameworks.

To complement quantitative analysis, thematic coding of five expert interviews was conducted. Interviewees included logistics platform architects, AI engineers, and blockchain strategists. Thematic analysis revealed four critical themes: automation confidence, regulatory ambiguity, system integration complexity, and data governance challenges.

Experts consistently emphasized improved trust in automated decision-making with AI-enhanced smart contracts. Where traditional contracts required manual inspection, AI-enabled logic facilitated autonomous triggering of shipment releases, dispute resolution, and dynamic pricing. AI-powered smart contracts operate on decentralized architectures, while many logistics companies still depend on centralized enterprise resource planning (ERP) systems. Experts reported high technical debt and conversion costs when synchronizing blockchain logs with legacy data lakes. Multiple participants flagged risks surrounding biased training data. Incomplete or skewed data from IoT trackers and warehouse logs distorted predictive model outputs. Furthermore, data latency and inconsistency between smart contract execution logs and physical shipment tracking interfered with model accuracy. A synthesis of both data streams underlines the transformative impact of AI in smart contract applications while also revealing operational and strategic frictions. The quantitative data affirms that increased AI levels correlate strongly with better logistics performance metrics. Qualitative narratives explain the technological, regulatory, and organizational barriers that inhibit the full potential of AI-smart contract integration. Regulatory bodies should expedite frameworks for AI accountability and blockchain traceability, as highlighted in the qualitative findings. Furthermore, platform developers must focus on hybrid architecture models capable of interacting with legacy ERP systems while maintaining blockchain-based auditability. Finally, increasing data quality via synchronized IoT, GPS, and RFID feeds will be pivotal to enhancing AI decision accuracy.

This study demonstrates notable trends, although certain limitations must be recognised. The transaction dataset, while varied, was confined to one blockchain platform. Expanded validation across several platforms may enhance external validity. The selection of interviewees was restricted to five specialists; broadening the panel geographically might provide more nuanced insights.

AI applications in logistics smart contracts provide clear advantages in transaction efficiency, cost reduction, and mistake minimisation. Quantitative investigation validated statistically

significant correlations between AI proficiency and operational success. Machine learning techniques, especially neural networks, demonstrated predictive efficacy in simulating logistical behaviour. Insights derived from expert interviews contextualised these findings, uncovering both strategic opportunities and systemic obstacles. The integration of AI and blockchain is transforming logistics automation. The efficacy of AI in autonomous smart contract execution relies on strong data pipelines, legal clarity, and system interoperability..

5. Conclusion

5.1. Conclusion

The integration of Artificial Intelligence (AI) in logistics smart contracts has demonstrated significant potential to enhance operational efficiency, reduce costs, and improve reliability in decentralized logistics platforms. This study, through a mixed-methods approach, combined the analysis of 112 blockchain-based smart contract transactions with insights from five expert interviews to offer a comprehensive understanding of AI's role in automated decision-making processes. Quantitative findings revealed that transactions with high levels of AI interaction exhibited faster execution times (mean = 13.5 seconds), lower transaction costs (mean gas fee =0.00185 ETH), and significantly fewer errors (error rate = 2.5%) compared to those with low AI involvement (mean execution time = 17.2 seconds, gas fee = 0.00221 ETH, error rate = 6.7%). Qualitative insights emphasized that despite these advantages, adoption challenges persist, especially regarding interoperability between AI modules and blockchain environments, algorithm transparency, and trust in AI-generated decisions. Experts noted that the success of AI integration depends heavily on technical literacy within logistics firms, robust governance frameworks, and the standardization of data input formats to facilitate seamless automation. The results affirm that AI-supported smart contracts are not only feasible but also beneficial in streamlining logistics workflows.

5.2. Policy implications

The findings of this study offer several implications for policymakers and industry leaders seeking to harness the advantages of AI in logistics smart contracts. *Firstly,* the observed performance improvements in transactions with higher AI involvement strongly support the development of incentive structures that promote AI integration within decentralized logistics platforms. *Secondly,* regulatory bodies should prioritize the establishment of AI and smart contract interoperability standards to ensure security, transparency, and accountability across logistics operations..

Furthermore, the disparity in performance outcomes across AI interaction levels suggests that strategic capacity-building initiatives are critical. Institutions could also support sandbox environments that allow experimentation with AI-enabled contracts before full-scale deployment. Lastly, government and industry consortiums must collaborate to establish blockchain-AI innovation hubs that foster experimentation and prototyping.

References

Agrawal, Tarun Kumar, Jannis Angelis, Wajid Ali Khilji, Ravi Kalaiarasan, and Magnus Wiktorsson, “Demonstration of a blockchain-based framework using smart contracts for supply chain collaboration,” *International Journal of Production Research*, 61 (2023), 1497–1516.

Alqarni, Mohammed Ali, Mohammed Saeed Alkatheiri, Sajjad Hussain Chauhdary, and Sajid Saleem, “Use of Blockchain-Based Smart Contracts in Logistics and Supply Chains,” *Electronics*, 12 (2023), 1340

Arumugam, S. S., Umashankar, V., Narendra, N. C., Badrinath, R., Mujumdar, A. P., Holler, J., and Hernandez, A. 2018. "*IoT Enabled Smart Logistics Using Smart Contracts*," in Proceedings of the 8th International Conference on Logistics, Informatics and Service Sciences (LISS), IEEE, pp. 1–6.

Bottoni, P., Gessa, N., Massa, G., Pareschi, R., Selim, H., and Arcuri, E. 2020. "*Intelligent Smart Contracts for Innovative Supply Chain Management,*" Frontiers in Blockchain (3), Article 535787.

Chukwu, Nnaji, Simo Yufenyuy, Eunice Ejiofor, Darlington Ekweli, Oluwadamilola Ogunleye, Tosin Clement, Callistus Obunadike, Sulaimon Adeniji, Emmanuel Elom, and Chinenye Obunadike, “Resilient Chain: AI-Enhanced Supply Chain Security and Efficiency Integration,” *International Journal of Scientific and Management Research*, 07 (2024), 46–65.

Groschopf, Wolfram, Mario Dobrovnik, and Christian Herneth, “Smart Contracts for Sustainable Supply Chain Management: Conceptual Frameworks for Supply Chain Maturity Evaluation and Smart Contract Sustainability Assessment,” *Frontiers in Blockchain*, 4 (2021), 506436.

Kumar, Ajay, Kumar Abhishek, Pranav Nerurkar, Muhammad Rukunuddin Ghalib, Achyut Shankar, and Xiaochun Cheng, “Secure smart contracts for cloud-based manufacturing using Ethereum blockchain,” *Transactions on Emerging Telecommunications Technologies*, 33 (2022), e4129.

Li, Hongzhi, Dezhi Han, and Mingdong Tang, “A Privacy-Preserving Storage Scheme for Logistics Data With Assistance of Blockchain,” *IEEE Internet of Things Journal*, 9 (2022), 4704–4720.

Lu, Yingfa, and Changbo Cheng, eds., “Front Matter: Volume 12260,” in *International Conference on Computer Application and Information Security (ICCAIS 2021)*, 2022 (Wuhan, China, SPIE).

Raja Santhi, Abirami, and Padmakumar Muthuswamy, “Influence of Blockchain Technology in Manufacturing Supply Chain and Logistics,” *Logistics*, 6 (2022), 15.

B1

Mapping the Techno-Functional Architecture of Green Manufacturing: A Patent-based Analysis

Yu-Chao Cheng, Cheng-Hsien Hsieh (Singapore University of Social Sciences)

The rapid expansion of e-commerce increasingly depends on intelligent manufacturing and logistics infrastructures capable of supporting data-driven and sustainable supply chains. While existing studies on artificial intelligence (AI) in e-commerce primarily focus on customer analytics, marketing automation, or platform algorithms, limited research has examined the technological architecture underpinning these systems. This study investigates the innovation landscape of sustainable manufacturing technologies that support digitally integrated commerce ecosystems through large-scale patent analytics. Using SciBERT-based semantic clustering, we analyse a portfolio of USPTO patents to identify the evolving technological structure of green manufacturing innovation. The findings indicate that green manufacturing innovation is not confined to generation technologies but is increasingly organised as an interconnected techno-functional stack combining digital intelligence, process chemistry, and thermal management. By mapping these technological layers, the study contributes a replicable data-driven framework for identifying technological convergence, emerging specialisations, and strategic inflection points relevant to the broader digital commerce ecosystem.

KEY WORDS

Green manufacturing, patent data, semantic clustering, artificial intelligence (AI), SciBERT

Mapping the Techno-Functional Architecture of Green Manufacturing: A Patent-based Analysis

Cheng, Y. C. and Hsieh, C. H.[*]
School of Business, Singapore University of Social Sciences
[*]*Corresponding author: chhsieh@suss.edu.sg*

Abstract

The rapid expansion of e-commerce increasingly depends on intelligent manufacturing and logistics infrastructures capable of supporting data-driven and sustainable supply chains. While existing studies on artificial intelligence (AI) in e-commerce primarily focus on customer analytics, marketing automation, or platform algorithms, limited research has examined the technological architecture underpinning these systems. This study investigates the innovation landscape of sustainable manufacturing technologies that support digitally integrated commerce ecosystems through large-scale patent analytics. Using SciBERT-based semantic clustering, we analyse a portfolio of USPTO patents to identify the evolving technological structure of green manufacturing innovation. The findings indicate that green manufacturing innovation is not confined to generation technologies but is increasingly organised as an interconnected techno-functional stack combining digital intelligence, process chemistry, and thermal management. By mapping these technological layers, the study contributes a replicable data-driven framework for identifying technological convergence, emerging specialisations, and strategic inflection points relevant to the broader digital commerce ecosystem.

Keywords: Green manufacturing, patent data, semantic clustering, artificial intelligence (AI), SciBERT

Introduction

Sustainability has increasingly become a central principle guiding global economic and policy development. Political and business leaders now widely recognise that environmental protection and economic growth are not mutually exclusive but mutually reinforcing pillars of sustainable development (Huesemann, 2003). The concept gained prominence following the publication of Our Common Future, produced by the World Commission on Environment and Development (1987), which defined sustainable development as meeting present needs without compromising the ability of future generations to meet their own. Within this framework, innovation and technological advancement play a critical role in enabling economic growth while reducing environmental pressures (Ghisetti et al., 2015). Increasing attention has therefore been directed towards sustainable production and consumption systems, particularly in sectors undergoing rapid digital transformation.

The expansion of the digital economy has significantly transformed patterns of production, consumption, and commercial interaction, with technological progress, widespread internet access, and smartphone adoption accelerating the rapid growth of e-commerce. As online transactions increasingly become embedded in global retail systems, the scale of e-commerce has expanded substantially. Strong digital infrastructure and high internet penetration have further stimulated online consumption, making e-commerce a central component of modern retail development (Liu et al., 2025). However, this growth also raises environmental concerns. Although e-commerce is not traditionally classified as a high-pollution industry, digital intelligence further enables firms to operationalise sustainability goals across complex production and logistics networks. Consequently, the convergence of digital commerce and industrial production necessitates a dual focus on green manufacturing and AI-driven logistics

The rise of e-commerce has reshaped retail structures and supply chain activities, while manufacturing remains a major contributor to resource consumption and environmental impacts. Consequently, integrating sustainability principles into both digital commerce and manufacturing processes has become increasingly important. Initiatives aligned with the United Nations Sustainable Development Goals emphasise responsible production, efficient resource utilisation, and clean energy transitions, highlighting the need for businesses to adopt sustainable operational models that balance economic efficiency with long-term environmental and social responsibility. Green manufacturing emphasises reducing environmental impacts through efficient resource utilisation, waste minimisation, and low-carbon production processes, while AI-driven supply chain optimisation employs machine learning, predictive analytics, and real-time data processing to improve decision-making across procurement, production, and distribution. This is especially pertinent within the sphere of climate change and the subsequent requirement for the large-scale expansion and dissemination of relevant technologies. To accurately measure the trajectory of these technological shifts and their impact on sustainability, empirical evidence from intellectual property (IP) filings becomes indispensable.

Patent data represents a critical instrument for both industrial and academic sectors, offering a robust, quantitative foundation for the analysis of technological progression, the diffusion of knowledge, and shifts in market competition (Cho et al., 2021). Furthermore, patent datasets facilitate the mapping of technological evolution over time, whilst offering granular details regarding nascent innovations and providing the necessary insights to identify historical, contemporary, and prospective technological trends (Altuntas et al., 2015). While traditional patent analysis provides a quantitative overview, this study incorporates a Large Language Model (LLM) to parse semantic nuances and identify latent themes that conventional metadata-based approaches often overlook. The rest of this paper is organised as follows. Section 2 provides briefing background for green manufacturing. Section 3 details the proposed methodology. Section 4 presents results and discussions, followed by summarising the key findings and industrial implications.

Green Manufacturing

Green manufacturing is primarily concerned with transforming corporate processes, manufacturing practices, and stakeholder perspectives to mitigate the industrial contributions to climate change and broader environmental degradation (Haleem et al., 2023). By adopting green and sustainable approaches, firms can enhance their operational effectiveness, competitiveness, and profitability. An increasing number of manufacturers recognise the diverse and enduring financial benefits associated with the integration of environmentally conscious practices (Leong et al., 2019). Achieving such outcomes, however, often necessitates the investigation, development, and deployment of intelligent technologies to identify and implement optimal practices that minimise environmental impact (Karuppiah et al., 2020). Defined as a collection of innovative production techniques, green manufacturing seeks to minimise adverse environmental effects through waste reduction, material recycling, and more efficient use of natural resources. Also referred to as cleaner production, this approach advocates for the modernisation of conventional manufacturing models in favour of sustainable systems that prioritise the prevention of waste over its subsequent management. In doing so, it expands the traditional "3R" framework, i.e., reduce, reuse, recycle, into a more comprehensive "5R" paradigm encompassing repair, reuse, refurbishment, rebuild, and recycling. It involves a fundamental change in corporate processes, manufacturing practices, and stakeholder mindsets to mitigate the industrial impact on climate change (Rehman and Seth, 2024).

The adoption of green manufacturing is increasingly critical as it enables firms to become more successful, competitive, and profitable. Beyond considerations of ethical responsibility, manufacturers are recognising both immediate and enduring financial advantages associated with the integration of environmentally conscious measures. Green manufacturing is essential for achieving business sustainability in an era characterised by greater consumer awareness regarding the environmental footprint of products. Given that manufacturing constitutes a fundamental driver of economic growth, transitioning towards sustainable practices, including the modernisation of production techniques in sectors such as casting and forging, is imperative for fostering sustainable development and addressing persistent economic constraints, notably low profit margins (Singh et al., 2018). Moreover, the

implementation of environmentally oriented policies assists firms in enhancing brand equity and reputation among ecologically mindful consumers, whilst simultaneously ensuring compliance with rigorous regulatory standards pertaining to air and water quality. In addition to the waste reduction, the transition towards renewable energy and biofuels constitutes a central strategy in reducing carbon emissions and operational costs within industrial systems. Facilities can generate electricity, heat, or fuel from renewable sources such as solar, wind, biomass, geothermal, hydropower, and waste-to-energy systems, thereby decreasing reliance on fossil fuels and enhancing long-term energy resilience. In parallel, the adoption of energy-efficient lighting, optimisation of lighting schedules in accordance with production planning, and routine equipment inspections contribute to sustained reductions in energy consumption and operating expenditure. Investments in energy-efficient machinery, alongside solar and wind installations, can substantially lower electricity demand over time. Operational measures including digitalisation and paperless processes further reduce material use and associated logistical costs (Rehman et al., 2014).

Advancements in sustainability require organisational commitment and collective participation. Internal communication that emphasises environmental responsibility fosters a culture of continuous improvement and strengthens employee engagement, reinforcing green practices across operations. More broadly, the integration of renewable energy, resource conservation techniques, and efficient production systems supports environmentally responsible manufacturing. In high-impact industries such as automotive manufacturing, process innovations including Just-in-Time production have contributed to reducing waste and improving resource efficiency, thereby mitigating environmental pressures associated with large-scale industrial activity (Zhang et al., 2019). Enhancing the efficiency of facilities and equipment represents a fundamental dimension of sustainable production. Emphasis is placed on reducing energy intensity through the adoption of high-efficiency technologies, including advanced electric motors, energy-saving machinery, and optimised lighting systems. Such measures not only decrease operational costs but also contribute to lower carbon emissions across production processes. Increasingly, digital monitoring systems enable firms to collect and analyse detailed data on energy consumption, water usage, carbon emissions, regulatory compliance, and incident management. Dedicated software platforms provide real-time environmental performance dashboards, allowing organisations to measure, evaluate, and improve sustainability metrics more effectively than traditional manual recording methods (Baumer-Cardoso et al., 2020). Green computing further strengthens facility-level sustainability. This approach promotes the environmentally responsible design, manufacture, use, and disposal of computing equipment, aiming to minimise lifecycle environmental impacts (Ashimaet al., 2021). Core strategies include improving the energy efficiency of processors, servers, and peripheral devices, reducing resource consumption, and addressing electronic waste. By lowering the carbon footprint associated with information technology systems, green computing supports broader sustainable development objectives while reinforcing operational efficiency within modern industrial environments.

Sustainable production necessitates a comprehensive assessment of a facility's social, environmental, and economic dimensions to ensure a harmonious relationship with the natural environment. Manufacturers adopting such environmentally conscious practices frequently position sustainability as a core component of the broader environmental strategy. This strategic orientation is partly driven by the recognition that maintaining consumer satisfaction is essential for long-term commercial success, as ecological preservation becomes an increasingly significant concern for many customers. Firms, consequently, are endeavouring to utilise resources in ways that safeguard the environment (Afum et al., 2020). Organisations at every stage of the value chain encounter challenges associated with the consumption of water, electricity, and chemical inputs. In contemporary civilisation, sustainability has emerged as a fundamental developmental objective for humanity. The social dimension of sustainability carries considerable societal implications, particularly as it directly affects human health, whilst also intersecting with issues such as labour costs and environmental degradation. Addressing the interconnected social, environmental, and economic challenges within the manufacturing sector requires careful examination of the linkages between these three perspectives. It is through the balanced development of society, the environment, and the economy that the manufacturing industry can be effectively guided towards a more sustainable trajectory (Sun et al., 2020).

Recent research indicates that AI technologies facilitate greener manufacturing by enabling accurate demand forecasting, optimised routing, intelligent inventory management, and dynamic production planning, thereby reducing energy consumption, material waste, and carbon emissions throughout the supply chain (Chen et al., 2024; Jiang, 2025). Furthermore, the integration of AI with Industry 4.0 technologies, such as the Internet of Things (IoT) and big-data analytics, supports the emergence of "green smart supply chains", where data-driven coordination across suppliers, manufacturers, and distributors enhances environmental performance and operational efficiency simultaneously (Zhang et al., 2025). Empirical studies also suggest that intelligent manufacturing systems can stimulate green innovation along supply chains through knowledge spillovers and improved coordination among supply chain partners (Mao and Shi, 2026). Consequently, AI-driven supply chain optimisation functions as a key enabler of green manufacturing, transforming sustainability from a compliance requirement into a strategically embedded capability within modern manufacturing ecosystems.

Research Methodology

3.1 Corpus Construction and Preprocessing

The empirical dataset consists of 5,095 patent abstracts retrieved from the United States Patent and Trademark Office (USPTO) database spanning 2005–2025. Documents were selected using a Boolean query combining sustainability and manufacturing-related terms ("energy-efficient" OR "resource-efficient" OR "sustainable manufacturing" OR "circular economy"). Duplicate records were removed using patent identifiers and abstract-level hash comparison to ensure corpus integrity.

Transformer-based models such as SciBERT have significantly improved the representation of domain-specific scientific language, enabling more accurate semantic analysis of technological knowledge structures (Burmaoglu et al., 2025). Text preprocessing was deliberately minimal to preserve contextual semantics required by transformer architectures. Abstracts were normalised via Unicode standardisation and lowercasing, followed by removal of extraneous non-alphanumeric characters and redundant whitespace. No stemming or lemmatisation was performed, as contextual language models encode morphological variations directly. Tokenisation was carried out using the SciBERT WordPiece tokeniser to ensure vocabulary alignment with the pretrained embedding space. Sequences exceeding the 512-token positional limit were truncated; empirical inspection indicated that only a small fraction of abstracts required truncation, thereby limiting semantic information loss.

3.2 Contextual Embedding Representation

To construct high-dimensional semantic representations, each abstract was encoded using the SciBERT model (scibert-scivocab-uncased), a transformer architecture pretrained on large-scale scientific corpora (Beltagy et al., 2019). Let an abstract consist of T tokens. The model produces contextual hidden states $H \in \mathbb{R}^{T\times 768}$, where 768 denotes the embedding dimension of the final transformer layer.

Document-level embeddings were computed using attention-masked mean pooling:

$$\mathbf{e} = \frac{1}{\sum_{i=1}^{T} m_i} \sum_{i=1}^{T} m_i \, H_i,$$

where m_i is the binary attention mask for valid tokens. Mean pooling was selected to mitigate positional bias and to capture distributed semantic content across patent abstracts, which typically allocate important information throughout the text rather than concentrating it at the beginning.

The resulting embedding matrix $E \in \mathbb{R}^{5095\times 768}$represents each abstract as a dense vector in continuous semantic space. Prior to clustering, embeddings were L_2-normalised to unit length. This transformation reduces vector magnitude variance and aligns Euclidean distance with cosine similarity geometry, which is better suited to contextual embedding spaces where angular separation carries greater semantic meaning than absolute scale differences.

3.3 Unsupervised Clustering Strategy

To identify latent thematic structure within the embedding space, two clustering paradigms were evaluated: HDBSCAN and k-means. HDBSCAN, a hierarchical density-based algorithm, was initially explored to detect clusters of arbitrary shape and varying density without requiring pre-specification of the number of clusters. However, when applied to the normalised 768-dimensional embedding manifold, HDBSCAN yielded a high proportion of samples classified as noise and produced clusters with substantial size imbalance. This outcome suggests relatively homogeneous point dispersion in the contextual embedding space, limiting the effectiveness of density contrast as a separation criterion. Although silhouette coefficients appeared numerically strong, inspection revealed that these values were inflated by the exclusion of boundary instances, thereby overstating practical separability.

Given these characteristics, a centroid-based partitioning approach was subsequently adopted (Beltagy et al., 2019). k-means clustering with $k = 5$, k-means++ initialisation, 20 random restarts, and a maximum of 300 iterations was implemented to minimise within-cluster sum of squared distances:

$$\sum_{j=1}^{k} \sum_{x_i \in C_j} \|x_i - \mu_j\|^2$$

where μ_j denotes the centroid of cluster C_j. The choice of $k = 5$ was guided by inspection of inertia curves and thematic interpretability of resulting partitions. Stability analysis across random seeds demonstrated high consistency in cluster assignments, indicating robust centroid convergence. Compared with density-based segmentation, k-means produced balanced and semantically coherent groupings, supporting clearer thematic delineation and downstream keyword extraction. Consequently, all reported clustering results are based on the k-means solution.

Analytical Results

To evaluate the structural distribution of the semantic clusters, the two-dimensional t-SNE projection of the 768-dimensional SciBERT embeddings is shown in Figure 1. Points are coloured according to their k-means cluster assignments, enabling qualitative assessment of thematic compactness and inter-cluster proximity. The visualisation reveals five distinguishable yet partially interconnected regions.

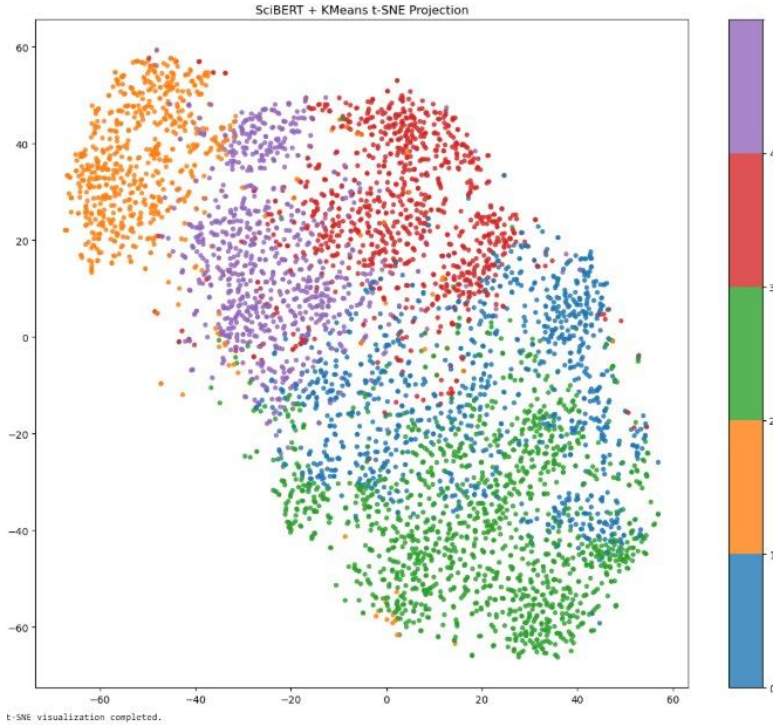


Figure 1. t-SNE visualisation of clustered patent based on SciBERT embeddings

Table 1 presents the five thematic clusters derived from SciBERT-based semantic grouping of 5,095 abstracts. The clusters differ both in scale and technological orientation. Cluster 2 (n = 1,506) is the largest group and centres on wireless communication, processors, network transmission, and computing architectures. The concentration of networking and memory-related terminology indicates a strong presence of digital coordination and data-processing technologies within the green energy domain. The size of this cluster suggests that computational optimisation and connectivity infrastructure constitute a substantial portion of recent patenting activity. The green cluster in Figure 1 occupies a sizeable and relatively central region in the projection. Its central positioning suggests connective relevance across multiple technological themes, consistent with the cross-cutting role of computational optimisation in modern energy systems.

Table 1 Top 10 Keywords per Cluster

Cluster ID	Number of Patents	Top 10 Keywords
0	1037	sensor, lighting, circuit, device, thermostat, efficiency, energy, generator, electrical, vehicle
1	695	distillation, process, catalyst, polymerization, solvent, plastics, polypropylene, reactor, ethanol, pyrolysis
2	1506	device, wireless, network, processor, computing, ethernet, transmitting, processing, efficiency, memory
3	963	lighting, heating, lamp, illumination, building, optical, thermal, light, cooling, insulation
4	894	cooling, heating, compressor, evaporator, drying, thermal, refrigerant, refrigeration, heater, heat

Cluster 0 (n = 1,037) captures sensor systems, circuit-level components, electrical devices, thermostatic control, and vehicle-related applications. Compared to Cluster 2's software-oriented profile, this group reflects hardware-level control and energy monitoring technologies. The co-occurrence of generator, electrical, and vehicle points to integration of energy devices with mobility and distributed power systems. The blue cluster in Figure 1 appears more dispersed and partially overlaps with Cluster 2. Containing sensors, circuit components, thermostatic controls, generators, and vehicle-related technologies, this cluster demonstrates broader semantic dispersion. Rather than indicating noise, the diffusion and overlap pattern implies integrative functionality, bridging electrical control mechanisms with digital networking infrastructures. The spatial adjacency between Clusters 0 and 2 highlights their joint contribution to sensing, actuation, communication, and orchestration across energy platforms. Cluster 0 is exemplified by an aircraft smart gauge that activates its analogue display only when sensor-detected changes exceed a threshold, conserving energy. By contrast, Cluster 2 features vRAN architectures that dynamically process codeblocks to optimise quality of service, enabling low-latency, energy-efficient edge networks.

Cluster 3 (n = 963) aggregates lighting, heating, insulation, optical, and building-related technologies. The vocabulary indicates energy performance improvements at the building and environmental interface, particularly in illumination efficiency and thermal regulation within constructed spaces. Cluster 4 (n = 894) is dominated by cooling, refrigeration, compressor, evaporator, and heat-exchange terminology. Unlike Cluster 3's architectural focus, this cluster concentrates on equipment-specific thermal control technologies, indicating continued innovation in mechanical energy-efficiency improvements.

Cluster 3 (the red cluster in Figure 1) and Cluster 4 (the purple cluster in Figure 1) occupy adjacent but separable regions of the embedding space. The red cluster shows moderate cohesion and centres on lighting, insulation, and building-related heating systems, indicating energy optimisation at the environmental or architectural scale. In contrast, the purple cluster appears slightly more compact and

gravitates toward refrigeration, compressors, evaporators, and heat-exchange components, reflecting equipment-level thermal regulation technologies. Their spatial proximity suggests conceptual linkage within broader energy efficiency research, while their separation indicates differentiation between building-scale and device-scale innovation domains. Cluster 3 comprises an electromotive drive for buoyant systems integrating a sealed electric motor with water-based cooling and streamlined design to enhance thermal regulation and efficiency. Cluster 4 comprises a modular semiconductor thermal management system using pump-driven fluid loops, heaters and refrigeration cycles for precise temperature control and adaptability.

Finally, Cluster 1 (n = 695) forms the most chemically oriented cluster, featuring distillation, catalysis, polymerisation, solvent use, and pyrolysis. This group reflects process engineering and material transformation technologies, including biofuel and polymer upgrading pathways. The brown cluster forms the most compact and internally cohesive grouping. Its dense spatial concentration reflects strong semantic consistency around biofuel production and chemical process engineering topics. The tight clustering suggests a well-defined technological theme with relatively limited conceptual dispersion, consistent with structured research trajectories in biomass conversion and material transformation processes. A chemical process converts sulfur-containing carbonaceous residues in an updraft gasifier via sub-stoichiometric oxidation with oxygen-containing gas and alkaline additives, producing sulfur-poor gaseous cleavage products that are subsequently oxidised to generate heat for producing heated aqueous media used in oil sands and oil shale bitumen extraction.

Across clusters, the distribution of abstracts reveals a technologically diversified innovation landscape. Digital and electronic coordination technologies (Clusters 2 and 0 combined) account for nearly half of the total abstracts, while thermal regulation, building efficiency, and chemical transformation represent complementary but distinct innovation trajectories. The clustering results therefore indicate that green energy patenting activity extends beyond generation technologies into process engineering, device optimisation, and digital system integration. Overall, the t-SNE visualisation indicates that green energy research is structured around a set of distinct yet interdependent technological domains. Biofuel process engineering forms a clearly bounded theme; thermal efficiency innovations bifurcate into environmental and equipment scales; and digital–electrical coordination technologies occupy a connective position across the embedding space. The clustering patterns therefore reveal not isolated research silos but a multi-domain innovation landscape in which process engineering, thermal management, and intelligent system integration co-evolve within the broader green energy ecosystem.

Conclusion

The LLM clustering findings strongly corroborate the theoretical foundations of green manufacturing described in the literature by empirically revealing the underlying techno-functional architecture of sustainability-oriented innovation. For instance, the literature highlights the necessity of intelligent technologies and digital monitoring for operational effectiveness. This is supported by Findings Cluster 2, the largest group, which focuses on computing architectures and wireless networks, alongside Cluster 0's emphasis on sensors and electrical coordination. Together, these clusters validate the literature's assertion that green smart supply chains and data-driven coordination are central to modern sustainability together form a digital-electrical integration layer. Additionally, the literature identifies biofuels and renewable energy as core strategies for reducing carbon emissions. This aligns with Cluster 1, which shows a compact, well-defined focus on biomass conversion and chemical process engineering. Furthermore, the literature's emphasis on green buildings and facility efficiency is echoed by Clusters 3 and 4, which detail innovations in lighting, insulation, and equipment-level thermal regulation, thereby articulating dual-layer structure of thermal efficiency innovation. Ultimately, the findings confirm that green manufacturing has evolved into an interdependent ecosystem where process engineering and intelligent system integration co-evolve to mitigate environmental degradation.

From a methodological perspective, this study further contributes a replicable LLM-driven patent analytics approach, demonstrating how semantic clustering can uncover latent technological structures that are not observable through conventional classification methods.

From an industrial perspective, these patterns indicate that digital system integration is no longer a supporting function but a core strategic capability. The convergence of artificial intelligence with Industry 4.0 technologies, including the IoT and big data analytics, is enabling the development of smart and environmentally responsive supply chains. Drawing upon advanced computational infrastructures and networked sensing systems, firms can implement AI-driven optimisation to enhance demand forecasting, inventory control, and adaptive production scheduling, thereby lowering energy intensity and material waste. The integration of hardware-based sensing with software-enabled coordination further facilitates real-time environmental performance management. In this way, sustainability is embedded within core operational strategy, strengthening long-term competitiveness while reducing ecological pressures.

References

Afum, E., Agyabeng-Mensah, Y., Sun, Z., Frimpong, B., Kusi, L. Y., and Acquah, I. S. K. 2020. "Exploring the link between green manufacturing, operational competitiveness, firm reputation and sustainable performance dimensions: a mediated approach," *Journal of Manufacturing Technology Management* (31:7), pp. 1417-1438.

Altuntas, S., Dereli, T., and Kusiak, A. 2015. "Forecasting technology success based on patent data," *Technological Forecasting and Social Change* (96), pp. 202-214.

Ashima, R., Haleem, A., Bahl, S., Javaid, M., Mahla, S. K., and Singh, S. 2021. "Automation and manufacturing of smart materials in additive manufacturing technologies using internet of things towards the adoption of industry 4.0," *Materials Today: Proceedings* (45:6), pp. 5081-5088.

Baumer-Cardoso, M. I., Campos, L. M., Santos, P. P. P., and Frazzon, E. M. 2020. "Simulation-based analysis of catalysers and trade-offs in Lean & Green manufacturing," *Journal of Cleaner Production* (242), 118411.

Beltagy I, Lo K, Cohan A. (2019). SciBERT: A pretrained language model for scientific text. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing Nov 3–7; Hong Kong, China. Association for Computational Linguistics; 2019. p. 3615–20.

Burmaoglu S, Aravacik ED, Soyler A. Divergent priorities, convergent goals: semantic analysis of publications and patents in healthcare smart contracts. Scientometrics 2025; 130:6579–612.

Chen, W., Men, Y., Fuster, N., Osorio, C., and Juan, A. A. 2024. "Artificial intelligence in logistics optimisation with sustainable criteria," Sustainability (16:21), 9145.

Cho, R. L-T., Liu, J. S., and Ho, M. H-C. 2021. "The development of autonomous driving technology: perspectives from patent citation analysis," *Transport Reviews* (41:5), pp. 685-711.

Ghisetti, C., Marzucchi, A., and Montresor, S. 2015. "The open eco-innovation mode. An empirical investigation of eleven European countries," *Research Policy* (44:5), pp. 1080-1093.

Haleem, A., Javaid, M., Singh, R. P., Suman, R., and Qadri, M. A. 2023. "A pervasive study on Green Manufacturing towards attaining sustainability," *Green Technologies and Sustainability* (1:2), 100018,

Huesemann, M. H. 2003. "The limits of technological solutions to sustainable development," Clean Technologies and Environmental Policy (5), pp. 21-34.

Jiang, Y. 2025. "Supply chain optimization and sustainability driven by artificial intelligence," Journal of Applied Economics and Policy Studies (18:6), 21-25.

Karuppiah, K., Sankaranarayanan, B., Ali, S. M., Chowdhury, P., and Paul, S. K. 2020. "An integrated approach to modeling the barriers in implementing green manufacturing practices in SMEs," Journal of Cleaner Production (265), 121737.

Leong, W. D., Lam, H. L., Ng, W. P. Q., Lim, C. H., Tan, C. P., and Ponnambalam, S. G. 2019. "Lean and green manufacturing—a review on its applications and impacts," *Process Integration and Optimization for Sustainability* (3), pp. 5-23.

Liu, X., Cui, W., and Zhang, S. 2025. "Better e-commerce less carbon emissions in China?," *Energy* (318), 134820.

Mao, Q., and Shi, B. 2026. "Intelligent manufacturing and green innovation along the supply chain," *Economic Modelling* (154), 107386.

Rehman, M. A., and Seth, D. 2024. Addressing transition conflict of green manufacturing to sustainable manufacturing: A drivers-based modeling strategy for improving business performance: An Indian experience," *Business Strategy and the Environment* (33:7), pp. 6387-6416.

Rehman, M. A., Shrivastava, R. R., and Shrivastava, R. L. 2014. "Evaluating green manufacturing drivers: an interpretive structural modelling approach," *International Journal of Productivity and Quality Management* (13:4), pp 471-494.

Singh, A., Philip, D., Ramkumar, J. and Das, M. 2018. "A simulation based approach to realize green factory from unit green manufacturing processes," Journal of Cleaner Production (182), pp. 67-81.

Sun, Y. Bi, K., and Yin, S. 2020. "Measuring and integrating risk management into green innovation practices for green manufacturing under the global value chain," *Sustainability* (12:2), 545.

World Commission on Environment and Development. 1987. *Our Common Future*. Oxford, UK: Oxford University Press.

Zhang, L., Innab, N., Shuhidan, S. M. et al. 2025. "Artificial intelligence-driven internet of things-based green supply chain for carbon reduction in sustainable manufacturing," *Journal of Environmental Management* (389), 126170.

Zhang, X. Ming, X., Liu, Z., Qu, Y., and Yin, D. 2019. "General reference model and overall frameworks for green manufacturing," *Journal of Cleaner Production* (237), 117757.

B1

Blockchain-enabled Scope 3 Disclosure Strategies in Platform Supply Chains.

Yuqian Wang (Renmin University of China)

Against rising consumer environmental awareness and regulation, platforms face strategic choices in carbon disclosure. Blockchain enables transparent Scope 3 emissions tracking but entails high costs and complex incentive coordination. This study constructs a three-tier Stackelberg game model including a supplier, platform, and 3PL firm, all engaged in emission reduction. Four blockchain-enabled strategies are compared: no disclosure, Scope 1–2 only, upstream Scope 3, and full-chain Scope 3. Results show that broader disclosure cuts overall supply chain emissions but causes internal carbon leakage as the platform free-rides on partners. Narrower disclosure better reduces emissions at moderate traceability costs, especially in high environmental-awareness markets. Low traceability costs encourage platforms to adopt broader disclosure, balancing ecology and economy, while suppliers and 3PL firms prefer free-riding under platform-only disclosure. Full-chain disclosure maximizes social welfare but hurts platform profits. The optimal subsidy is negatively related to consumer environmental awareness, and full-chain disclosure needs more subsidies than upstream disclosure at low costs. This study informs platforms and policymakers on blockchain-driven disclosure and incentives.

KEY WORDS

Blockchain, Carbon disclosure, Scope 3 emissions, Platform supply chain, Incentive design, Carbon leakage

Blockchain-enabled Scope 3 Disclosure Strategies in Platform Supply Chains

Wang, Y. Q.[a]

[a] *Business Department, Renmin University of China*

E-mail: wangyuqian01@ruc.edu.cn

Abstract

Against rising consumer environmental awareness and regulation, platforms face strategic choices in carbon disclosure. Blockchain enables transparent Scope 3 emissions tracking but entails high costs and complex incentive coordination. This study constructs a three-tier Stackelberg game model including a supplier, platform, and 3PL firm, all engaged in emission reduction. Four blockchain-enabled strategies are compared: no disclosure, Scope 1–2 only, upstream Scope 3, and full-chain Scope 3. Results show that broader disclosure cuts overall supply chain emissions but causes internal carbon leakage as the platform free-rides on partners. Narrower disclosure better reduces emissions at moderate traceability costs, especially in high environmental-awareness markets. Low traceability costs encourage platforms to adopt broader disclosure, balancing ecology and economy, while suppliers and 3PL firms prefer free-riding under platform-only disclosure. Full-chain disclosure maximizes social welfare but hurts platform profits. The optimal subsidy is negatively related to consumer environmental awareness, and full-chain disclosure needs more subsidies than upstream disclosure at low costs. This study informs platforms and policymakers on blockchain-driven disclosure and incentives.

Keywords: Blockchain, carbon disclosure, Scope 3 emissions, platform supply chain, incentive design, carbon leakage

Introduction

Climate change, driven by rising carbon emissions, is one of the 21st century's most urgent challenges (Downar, et al., 2021). Platform-based firms are significant contributors, and eco-conscious consumers increasingly prioritize ESG factors, with over 70% willing to pay more for sustainable products (PwC, 2023). In response, platforms like Amazon, Meituan, and JD Logistics are integrating carbon reduction into their strategies to gain competitive advantages (Tian, and Bai, 2025).

Carbon disclosure regulations incentivize platforms to reduce emissions. Per the GHG Protocol, emissions are categorized into three scopes: Scope 1 (direct), Scope 2 (indirect energy-related), and Scope 3 (full value chain indirect emissions). Notably, Scope 3 accounts for 75% of average US corporate emissions, and CDP 2021 data shows supply chain emissions are 11.4 times higher than direct operational emissions (Hettler, and Graf-Vlachy, 2023). However, Scope 3 disclosure faces challenges: lack of standardized indicators, inefficient traditional data transmission, data privacy concerns, and consumer skepticism about greenwashing (over 50% of European consumers struggle to distinguish genuine eco-claims (Tian, and Bai, 2025)).

Blockchain, with its immutability and transparency, addresses these issues by enabling secure supply chain emissions tracking. It supports Scope 3 disclosure through multi-party data sharing, smart contract-based incentives, and enhanced consumer trust (Cui, Y., et al., 2024; Cui, Y., et al., 2023; Guo, et al., 2023; IBM, 2017; Kaplan, and Ramanna, 2021). We identify four platform disclosure strategies: no disclosure (Strategy N), Scope 1-2 only (Strategy S), upstream Scope 3 (Strategy U), and full-chain Scope 3 (Strategy C), with practical cases shown in Table 1. Blockchain implementation incurs

traceability costs (data storage, analysis) borne by policymakers and platforms (Pun, et al., 2021; Wu, J. H., and Wang, X., 2023; Wu, J. H., and Zhao, C. C., 2024), supported by a five-tier framework (Data Collection to Visualization Layers) ensuring efficient, tamper-proof data management.

Despite growing research on blockchain and emission reduction, its application in Scope 3 disclosure remains underexplored. This study addresses three key questions: (1) Does broader disclosure ensure greater emissions reduction? (2) How do platform disclosure strategies impact supply chain members, consumer surplus, and social welfare? (3) How to design optimal incentive subsidies?

We develop a game-theoretic model focusing on a platform-centric three-tier supply chain. Key findings show blockchain-supported disclosure effectiveness depends on traceability costs and market factors: broader disclosure reduces overall emissions but causes carbon leakage; moderate traceability costs favor narrower disclosure (especially in high environmental-awareness markets); lower costs support broader disclosure, though full-chain disclosure (Strategy C) maximizes social welfare but conflicts with platform profits; optimal subsidies correlate inversely with consumer environmental awareness.

This research contributes by modeling a platform-centric three-tier supply chain, integrating blockchain features with tiered disclosure, and designing incentive mechanisms. Subsequent sections include literature review, model setup, analysis, extensions, and conclusions.

Literature Review

This study is related to three streams of literature: low-carbon supply chain management, blockchain adoption in carbon emission reduction and carbon emissions disclosure.

Low-carbon Supply Chain Management

Low-carbon supply chain management has emerged as a key research focus. Firstly, a body of research has addressed pricing strategies and carbon policy design for low-carbon supply chains. Fan, et al., (2023) compare the effects of emission caps and carbon pricing on corporate investment and production, revealing the critical influence of policy uncertainty. Secondly, extensive research has examined incentive mechanisms for emission reduction. Dou, et al., (2024) propose a collaborative model where suppliers transfer surplus emission quotas to manufacturers and jointly invest in reduction projects, examining its impact on low-carbon performance and policy effectiveness from a dynamic perspective. Gopalakrishnan, et al., (2021) draws on cooperative game theory to develop a carbon footprint balancing mechanism for supply chains using either carbon taxes or internal pricing, which enables the equitable allocation of the total emissions burden among the participating firms. On the other hand, some scholars have investigated how government regulation and environmental awareness influence corporate decision-making. Distinct from existing research focusing on two-tier supply chains and binary disclosure choices, this study examines three-tier platform supply chains where all three members undertake emission reduction efforts. Furthermore, we introduce blockchain-based disclosure strategies to analyze the trade-offs between members' emission reduction efforts and economic consideration.

Blockchain Adoption in Carbon Emission Reduction

Extensive researches have focused on blockchain technology characteristics enabling sustainable supply chains, investigating the impact on supply chain carbon reduction. A stream of research has examined the dual function of blockchain technology in achieving concurrent economic benefits (e.g., brand premiums, profitability (Wu, J. H., Cheng, et al., 2024; Yu, et al., 2025)) and environmental benefits (e.g., carbon emission reduction, green production (Zhang, Y., et al., 2024; Zhuo, et al., 2025)) within supply chains. Yu, et al., (2025) develop a dynamic game model revealing that blockchain technology enhances low-carbon e-commerce supply chain performance through brand effects in monopoly markets, while its effectiveness in oligopoly markets depends critically on competition intensity and brand loyalty enhancement. Existing literature overlooks the distinctiveness of Scope 3 disclosures and overgeneralizes the benefits of blockchain technology. This study integrates

blockchain's threefold characteristics—emission reduction efforts through traceability, enhanced market demand via transparency, and automated execution through smart contracts. It addresses the previously neglected issues of free-riding behavior and the incentive automation mechanisms driven by smart contracts.

Carbon Emissions Disclosure

The stream of operations management literature focusing on carbon emissions disclosure is more closely related to our work. Chen, Y., et al., (2025) propose that blockchain enhances coordination in carbon emissions data collection among supply chain members, thereby facilitating corporate Scope 3 carbon disclosure, revealing blockchain's role in promoting Scope 3 disclosure. Zhao, et al., (2026) analyzed the efficiency of two carbon emission reduction regulatory frameworks including the Indirect Source Rule (ISR) and the Direct Source Rule (DSR). Shi, et al., (2025) empirically demonstrate that while mandatory disclosure policies generally reduce total supply chain emissions, voluntary disclosure mechanisms may lead to an increase in Scope 3 emissions. From the perspective of a purchasing firm, Kalkanci, and Plambeck, (2020) comparatively analyze the decision-making logic for enterprises to understand and reduce their suppliers' carbon emissions under voluntary versus mandatory disclosure scenarios targeting investors. Our paper contributes to this literature by analyzing how scope expansion impacts profits and emission reduction incentives, and resolves the misalignment between platform profit and social welfare with targeted strategy interventions.

Model

Consider a three-tier supply chain comprising a supplier, a platform, and a third-party logistics provider. The platform sells products to consumers at a retail price p, incurring a wholesale price w per unit sold to the supplier and a transportation fee t per unit to the logistics provider. In subsequent context, the subscript "s" denotes the supplier, "p" denotes the platform, and "l" denotes the third-party logistics provider. The superscripts "N", "U", "S", and "C" represent Strategy N, Strategy U, Strategy S, and Strategy C respectively. Unlike studies that focus solely on single-or dual-layer carbon reduction efforts (Gopalakrishnan, et al., 2021; He, et al., 2025; Zhang, Y., et al., 2024; Wang, Y., et al., 2021), we contend that all three parties within the platform supply chain generate emissions and must reduce their carbon footprint. Consistent with this observation, we assume an additive form for the CFR efforts of the three parties in the demand function, which is consistent with operation literature (Ardia, et al., 2023; Tian, and Bai, 2025). Additionally, products have an initial carbon footprint e0, and the overall carbon footprint of the supply chain is jointly reduced through the emission reduction behaviors of various entities.

However, in case without blockchain-enabled disclosure, limited traceability and transparency can undermine the trust of CFR claims, as noted in sustainability reports (e.g., Wim Vermeulen 32). Let ρ ($0 < \rho < 1$) denote the consumers trust in CFR level without blockchain-enabled disclosure, where $\rho < 1$ reflects the mistrust regarding CFR claims, thereby reducing the willingness to pay. We assume that E_t^j represents the consumers perceived carbon emission $E_t^j = e_0 - \sum_{i=p,s,l}^{i} E_i^j \; j \in \{N, U, S, C\}$, if the CFR level of member i is disclosed under Strategy j, then $E_i^j = e_i^j$, if not disclosed, then $E_i^j = \rho e_i^j$. We define the market demand function as $d = a - p_0 - hE_t^j$, where a represents the market basic demand. Consumers environmental awareness is represented by parameter h ($0 < h < 1$), indicating that eco-conscious consumers are willing to pay a premium. The platform, supplier and the third-party logistic provider reduce emissions through DER efforts and emissions trading. The CFR cost is $k\frac{e_i^2}{2}$ $i \in \{p, s, l\}$, where k represents the CFR effort cost coefficient. The cost structure is consistent with the prior literature [10]. Blockchain technology plays a dual role. On the one hand, it eliminates consumer mistrust about carbon disclosure information through its tamper-proof feature to ensure data authenticity (Wu, J. H., Cheng, et al., 2024; Zhuo, et al., 2025; Zhang, Y., et al., 2024). On the other hand, the platform provides incentives to members participating in disclosure via smart contracts, awarding γ per unit of effort expended. Given these two functions, the blockchain-driven carbon disclosure strategy incurs a unit cost of c_b for traceability at each tier (Cui, Y., et al., 2024; Cui, Y., et

al., 2023; Zhang, K., et al., 2024), the unit cost is several times the c_b for each layer disclosed. For ease of reference, we provide a summary of notations in Table 1.

Table 1. Summary of Notations

Notation	Description
h	Consumers' environmental awareness.
ρ	Consumer trust on CFR levels.
e_0	The product's initial carbon footprint.
c_b	Blockchain-usage cost per unit product for a single disclosure tier.
m	Commission rate.
η	The blockchain-driven efficiency improvement in CFR.
i	Fixed cost coefficient.
k	CFR effort cost coefficient.
ω	Wholesale price for the product.
p_0	Retail price for the product.
e_p	Platform's unit product CFR level.
e_l	Logistic holder's unit product CFR level.
e_s	Supplier's s unit product CFR level.
d	Market demand.
π_i^j	Profit for i in case j. $i \in \{p, s, l\}$, $j \in \{N, B\}$.

The following section outlines the logical rationale, demand implications, and profit functions associated with each disclosure strategy. A schematic representation of these strategies is presented in Figure 1.

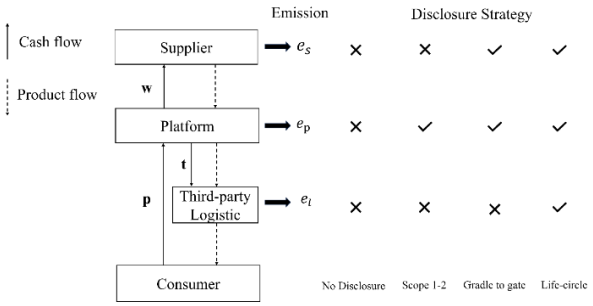


Figure 1. Model Illustration

No disclosure (Strategy N)

The platform does not disclose any carbon emission information. Market demand is only regulated by the initial carbon footprint and consumer trust, with the demand function $d^N = a - p_0 - h(e_0 - \rho e_s - \rho e_l - \rho e_p)$. And the profit function of the platform, the supplier and the logistic provider are $\pi_p^N = (p_0 - w - t)d^N - k\frac{e_p^2}{2}$, $\pi_s^N = wd^N - k\frac{e_s^2}{2}$ and $\pi_l^N = td^N - k\frac{e_l^2}{2}$.

Scope 1-2 Disclosure (Strategy S)

The platform only discloses its own direct emissions (Scope 1) and indirect emissions from purchased energy (Scope 2). The blockchain usage cost is $c_b$per unit, and the consumer trust in the platform's emission reduction level is increased to 1, with the demand function $d^S = a - p_0 - h(e_0 - \rho e_s - \rho e_l - e_p)$. In this case, the profit function of the platform, the supplier and the logistic provider are $\pi_p^S = (p_0 - w - t - c_b)d - k\frac{e_p^2}{2}$, $\pi_s^S = wd - k\frac{e_s^2}{2}$ and $\pi_l^S = td - k\frac{e_l^2}{2}$.

Upstream Scope 1-3 Emission Disclosure (Strategy U)

The platform discloses its own emissions and the upstream supplier's emissions. The blockchain usage cost increases to $2c_b$, the consumer trust in the supplier CFR level is increased to 1, and the platform provides incentives subsidy γ according to CFR level of the supplier, with the demand function $d^U = a - p_0 - h(e_0 - e_s - \rho e_l - e_p)$. In this case, the profit function of the platform, the

supplier and the logistic provider are $\pi_p^U = (p_0 - w - t - 2c_b)d - k\frac{e_p^2}{2} - e_s$, $\pi_s^U = wd - k\frac{e_s^2}{2} + e_s\gamma$ and $\pi_l^U = td - k\frac{e_l^2}{2}$.

Closed-Loop Scope 1-3 Emission Disclosure (Strategy C)

When the full-chain (supplier-platform-logistics provider) emission information is disclosed, the blockchain usage cost increases to $3c_b$, the consumer trust in all entities' CFR levels is increased to 1, and the platform provides incentives subsidy γ to both the supplier and logistic provider, with the demand function $d^C = 1 - p_0 - h(e_0 - e_s - e_l - e_p)$. In this case, the profit function of the platform, the supplier and the logistic provider are $\pi_p^C = (p_0 - w - t - 3c_b)d - k\frac{e_p^2}{2} - (e_l\gamma + e_s\gamma)$, $\pi_s^C = wd - k\frac{e_s^2}{2} + e_s\gamma$ and $\pi_l^C = td - k\frac{e_l^2}{2} + e_l\gamma$.

The sequence of the game is as follows: the supplier, the platform and the logistic provider decide CFR level e_p, e_s and e_l. In the following analysis, four disclosure strategies are represented by Strategy N, Strategy S, Strategy U, and Strategy C respectively.

Comparison and Analysis

CFR Performance Under Different Disclosure Strategy

The following propositions in this section examine the environmental impacts of different strategies according to each supply chain member and the entire supply chain.

Proposition 1 *(i)For the platform:* $e_p^C < e_p^U < e_p^S$; $e_p^S > e_p^N$ *when* $c_b < \widehat{c_1} = (p - t - w)(1 - \rho)$; $e_p^U > e_p^N$ *when* $c_b < \widehat{c_2} = \frac{1}{2}(p - t - w)(1 - \rho)$; $e_p^C > e_p^N$ *when* $c_b < \widehat{c_3} = \frac{1}{3}(p - t - w)(1 - \rho)$, *where* $\widehat{c_1} > \widehat{c_2} > \widehat{c_3}$. *(ii)For the supplier:* $e_s^C = e_s^U > e_s^S = e_p^N$. *(iii) For the logistics provider:* $e_l^C > e_l^U = e_l^S = e_l^N$.

Proposition 1 compares the CFR levels of the platform, the supplier and the third-party logistics provider under different strategies, revealing a crucial conclusion: a higher tier of disclosure strategy would lead to *carbon leakage* within the supply chain. It demonstrates that the sensitivity of platform CFR level to traceability costs increases significantly as the disclosure tier expands. Platforms exhibit "*emissions reduction free-riding*" behavior: its own CFR effort decreases strictly as disclosure tiers increase. This occurs because the platform bear rising traceability costs, and upstream or downstream entities' CFR efforts substitute for the platform's own motivation to reduce emissions. Consistent conclusions are supported by empirical research [41]. Blockchain-driven upstream disclosure Strategy U or C eliminates consumer trust discounts ρ and exert incentive subsidy γ, enabling a qualitative leap in the supplier emission reductions. Under both Strategy U and C, the supplier receives platform incentive subsidy, resulting in identical emission reduction levels. The logistics provider receives platform incentive subsidy γ only under Strategy C. Consequently, her emission reductions remain indistinguishable.

Proposition 2 (CFR Performance of the Total Supply Chain) *(i)* $e_t^C > \{e_t^S, e_t^U, e_t^N\}$ *when* $c_b < \widetilde{c_1}$; $e_t^U > \{e_t^S, e_t^C, e_t^N\}$ *when* $\widetilde{c_1} < c_b < \widetilde{c_2}$; $e_t^S > \{e_t^C, e_t^U, e_t^N\}$ *when* $\widetilde{c_2} < c_b < \widetilde{c_3}$. *(ii)* $\frac{\partial \widetilde{c_1}}{h} < 0$, $\frac{\partial \widetilde{c_2}}{h} < 0$; *(iii)* $\frac{\partial \widetilde{c_n}}{\rho} < 0, n \in \{1,2,3\}$

Proposition 2 *(i)* demonstrates that the optimal carbon disclosure strategy for environment performance is fundamentally cost-dependent. When the unit traceability costs exceed another critical threshold (we refer to as the "emission reduction traceability costs threshold"), the optimal strategy for maximizing emission reduction sequentially degrades from Strategy C to Strategy U, then to Strategy S, and ultimately to Strategy N, as illustrated in Figure 2. In this process, the platform must balance the inhibitory effect of expanding disclosure scope on its own emission reduction effort against the

promotive effect on upstream and downstream partners' abatement. If traceability costs are excessively high, the inhibitory effect dominates, making it optimal to reduce the disclosure tier. Conversely, in a low-cost environment, Strategy C is more feasible.

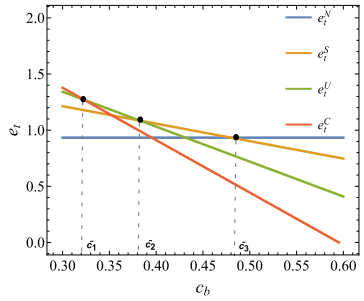


Figure 2. Impact of c_b on Total CFR level

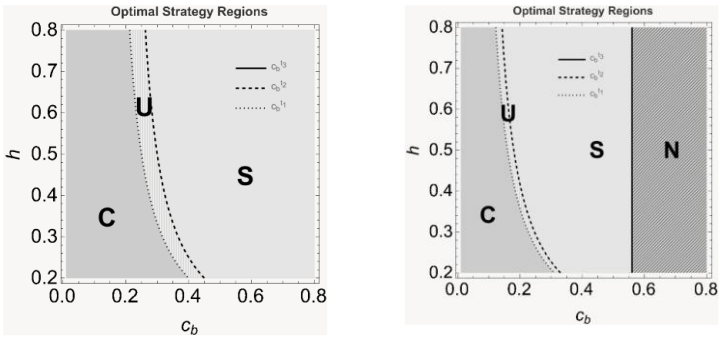


Figure 3. Platform's Disclosure Strategy Selection Based on total CFR Level ($e_0 = 2, a = 5, w = 0.4, t = 0.3, p = 3.5, \rho = 0.5, k = 0.3, \eta = 0, \gamma = 0.05$)

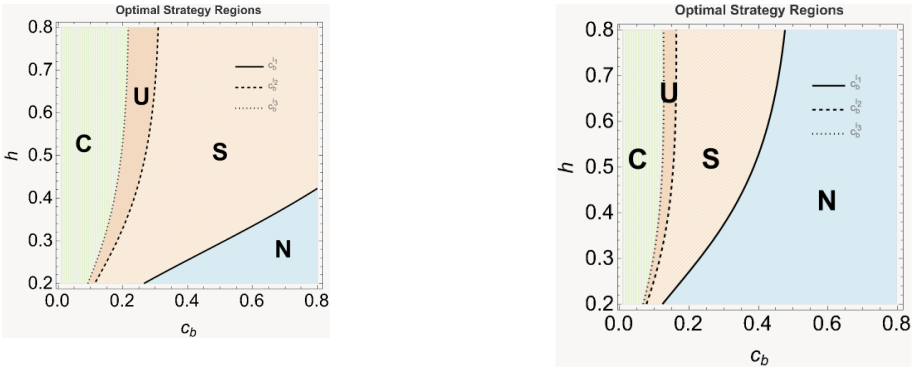
Additionally, Proposition 2 *(ii)* and Figure 6 present that the "emission reduction traceability costs threshold" decreases as consumer environmental awareness h increases.Instead, a narrower disclosure scope centralizes the platform's control over the emission reduction process, thereby strengthening its willingness to reduce emissions. For example, in the early stages of the electric vehicle industry when consumer environmental awareness was relatively low, Tesla differentiated itself by actively disclosing supply chain emissions (Strategy C). Proposition 2 *(iii)* implies that consumer trust decreases the "emission reduction traceability cost threshold," as visualized by comparing the left panel ($\rho = 0.5$) and the right panel ($\rho = 0.8$) of Figure 3. This indicates that low market trust enhances the value of blockchain-enabled disclosures, which amplify the incentive effect of emission reduction and increase the platform's acceptance to traceability costs.

Supply chain member profitability under different disclosure strategy

To provide a clearer explanation of the adoption of different strategies and their impact on supply chain members, the following comparative analysis examines the profitability of the platform, the supplier, third-party logistics provider, and the supply chain as a whole. Based on this, specific recommendations are proposed for the adoption of different disclosure strategies.

Proposition 3 (Optimal Strategy given Platform Profit) The platform's disclosure adoption strategy is: *(i)*. When $0 < c_b < c_b^{p_1}$, the platform adopts strategy C; *(ii)*. When $c_b^{p_1} < c_b < c_b^{p_2}$, the platform adopts strategy U; *(iii)*. When $c_b^{p_2} < c_b < c_b^{p_3}$, the platform adopts strategy S; *(iv)*. When $c_b^{p_3} < c_b < c_b^{p_{max}}$, the platform adopts strategy N. $c_b^{p_1}$, $c_b^{p_2}$ and $c_b^{p_3}$ can be seen in the Appendix B.

Proposition 3 points out that the platform's optimal disclosure strategy is determined by traceability costs. Lower traceability costs enable platforms to expand the coverage of carbon disclosure. Notably, there exist three traceability cost thresholds (referred to as the "strategy traceability costs threshold") below which the disclosure tier decreases, and these thresholds increase with the improvement of consumers' environmental awareness. The core logic is that profit-oriented strategy selection partially overlaps with emission reduction-oriented selection but has different priorities, which is strongly supported by the comparative analysis of the two panels in Figure 7. It shows the strategy distribution under platform profit maximization: at low c_b, the platform prioritizes Strategy C because the demand growth brought by full-chain disclosure (derived from consumers' environment preferences) exceeds the traceability costs. At moderate c_b, the U strategy balances the demand premium from upstream disclosure and moderate costs. At high c_b, the Strategy S transmits partial green signals at low costs, avoiding demand loss caused by the Strategy N. Figure 8 further supplements the win-win situation, benefiting both the profit and the emission reduction. When the platform considers both profit and emission reduction, the optimal region for Strategy C only exists in the "low cost" scenario



(a) $\rho = 0.5$ (b) $\rho = 0.8$

Figure 4. Platform's Disclosure Strategy Selection Based on Platform's Profitability ($\mathbf{e_0 = 2, a = 5, w = 0.4, t = 0.3, p = 3.5, k = 0.3, \gamma = 0.05}$)

remain constant, the higher the consumers' environmental awareness, the more optimal it is to adopt a disclosure strategy with more tiers.

Conclusion

This study develops a game-theoretic model of a three-tier supply chain comprising a supplier, a platform, and a third-party logistics provider to examine the implications of blockchain-enabled carbon disclosure strategies. Three main findings emerge.

First, broader disclosure does not necessarily yield superior emission reductions. While full-chain disclosure achieves optimal supply-chain-wide abatement when traceability costs fall below a critical threshold, it paradoxically enables the platform to free-ride on partners' efforts, inducing internal carbon leakage and creating misalignment between firm-level and supply-chain-level reductions. As disclosure scope widens, the platform's own abatement motivation may weaken due to rising cost burdens. Notably, in markets with high consumer environmental awareness, narrowing the disclosure scope helps the platform retain control over abatement processes and strengthens its reduction incentives. Second, optimal strategy selection depends jointly on traceability costs, consumer environmental awareness, trust levels, and incentive subsidy.

Several important managerial implications emerge from the analysis. For platforms, it is essential to balance disclosure breadth with technological costs. A phased approach is advisable: start with platform-only disclosure to build green brand equity, then gradually expand to upstream or full-chain disclosure as traceability costs decline through technological advances or cost-sharing mechanisms. To align environmental and economic outcomes, platforms should tailor their strategies to industry-specific levels of consumer environmental awareness and trust, and design incentives accordingly. Second, while we consider three disclosure tiers, real-world implementations often involve hybrid or partial disclosure—warranting further modeling. Third, the role of regulatory interventions, such as carbon taxes or mandatory disclosure laws, could be integrated to assess policy–technology interactions. Finally, empirical validation of the proposed model using real blockchain deployment data would strengthen the practical relevance of our findings.

References

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. 2023. "Climate Change Concerns and the Performance of Green vs. Brown Stocks," *Management Science* (69:12), pp. 7607-7632.

Chen, S., Zhang, Q. Q., and Zhou, Y. P. 2019. "Impact of Supply Chain Transparency on Sustainability under NGO Scrutiny," *Production and Operations Management* (28:12), pp. 3002-3022.

Chen, Y., Feng, Y., Lai, K.-H., and Zhu, Q. 2025. "Value of blockchain for scope 3 carbon disclosure: The moderating role of data processing technologies," *International Journal of Production Economics*, doi: 10.1016/j.ijpe.2024.109445.

Cui, Y., Gaur, V., and Liu, J. C. 2024. "Supply Chain Transparency and Blockchain Design," *Management Science* (70:5).

Cui, Y., Hu, M., and Liu, J. C. 2023. "Value and Design of Traceability-Driven Blockchains," *Manufacturing & Service Operations Management* (25:3), pp. 1099-1116.

Dou, L., Dai, Y., Song, H., Shu, Y., and Li, H. 2024. "Supply chain low-carbon cooperation under the partially paid carbon allowances policy: a dynamic perspective," *International Journal of Production Research*, Early Access, doi: 10.1080/00207543.2024.2444543.

Downar, B., Ernstberger, J., Reichelstein, S., Schwenen, S., and Zaklan, A. 2021. "The impact of carbon disclosure mandates on emissions and financial operating performance," *Review of Accounting Studies* (26:3), pp. 1137-1175.

Fan, X., Chen, K., and Chen, Y.-J. 2023. "Is Price Commitment a Better Solution to Control Carbon Emissions and Promote Technology Investment?," *Management Science* (69:1), pp. 325-341.

Gao, F., and Souza, G. C. 2022. "Carbon Offsetting with Eco-Conscious Consumers," *Management Science* (68:11), pp. 7879-7897.

Gopalakrishnan, S., Granot, D., Granot, F., Sosic, G., and Cui, H. 2021. "Incentives and Emission Responsibility Allocation in Supply Chains," *Management Science* (67:7), pp. 4172-4190.

Guo, S., Sun, X., and Lam, H. K. S. 2023. "Applications of Blockchain Technology in Sustainable Fashion Supply Chains: Operational Transparency and Environmental Efforts," *IEEE Transactions on Engineering Management* (70:4), pp. 1312-1328.

He, L., Han, R., Hou, L., and Yue, X. 2025. "Product Line Design and Channel Configuration in Low-Carbon Supply Chains," *IEEE Transactions on Engineering Management* (72), pp. 78-95.

Jiang, F., Huang, W., Yang, J., and Duan, H. 2025. "Retailer involvement in eco-conscious consumer-oriented carbon footprint reduction," *European Journal of Operational Research* (322:3), pp. 795-811.

Shi, Y. L., Tang, C. S., and Wu, J. 2025. "Are Firms Voluntarily Disclosing Emissions Greener?," *Production and Operations Management* (34:8), pp. 2363-2377.

Song, S. N., Dong, Y., Kull, T., Carter, C., and Xu, K. F. 2023. "Supply chain leakage of greenhouse gas emissions and supplier innovation," *Production and Operations Management* (32:3), pp. 882-903.

Tian, Y., and Bai, C. 2025. "Blockchain Adoption in the Low-Carbon Supply Chain: A Comparative Analysis of Different Adoption Modes," *IEEE Transactions on Engineering Management* (72), pp. 1559-1575.

Wang, Y., Yu, Z., Jin, M., and Mao, J. 2021. "Decisions and coordination of retailer-led low-carbon supply chain under altruistic preference," *European Journal of Operational Research* (293:3), pp. 910-925.

Wu, J. H., Cheng, H. Y., and Li, L. Y. 2024. "Blockchain-enabled downward traceability to combat gray market," *International Journal of Production Economics* (278), Art no. 109415.

Wu, J. H., and Wang, X. 2023. "Platform-leading blockchain adoption for traceability under upstream competition," *Annals of Operations Research* (348:3), pp. 2107-2143.

Wu, J. H., Yao, J. Z., and Wang, Y. Q. 2025. "Platform's role in green vs. non-green product competition: Blockchain transparency and selling contract selection," *International Journal of Production Economics* (288), Art no. 109703.

Wu, J. H., and Zhao, C. C. 2024. "Blockchain-Empowered Market Entry Strategy for a Green Product," *IEEE Transactions on Engineering Management* (71), pp. 8886-8900.

Yu, L., Yang, M., and Zhang, Z. 2025. "Introducing blockchain into low-carbon e-commerce supply chains: a differential game analysis in monopoly and oligopoly markets," *International Journal of Production Research* (63:14), pp. 5263-5291.

Zhang, K., Choi, T. M., Chung, S. H., Dai, Y., and Wen, X. 2024. "Blockchain adoption in retail operations: Stablecoins and traceability," *European Journal of Operational Research* (315:1), pp. 147-160.

Zhang, Y., Zhang, J., Zhou, Y., Zhao, H., and Cheng, Y. 2024. "Blockchain adoption and mode selection strategies for remanufacturing supply chain under cap-and-trade policy," *Computers & Industrial Engineering* (192), Art no. 110246.

Zhao, S., Gui, L. Y., and Cui, S. L. 2026. "Emission Reduction Through Regulating Indirect Sources," *Manufacturing & Service Operations Management* (28:1).

Zhuo, X. P., Sun, Y. J., and Zhou, S. R. 2025. "Impact of Consumers Traceability Awareness on Blockchain Adoption in Supply Chains," *IEEE Transactions on Engineering Management* (72), pp. 1140-1153.

B1

From Engagement to Collapse: A Dynamic Change-Point Perspective on Gamified OMO Customer Experience.

Su Jun Zhu, ing Zhang (KOOKMIN UNIVERSITY)

Gamified Online-Merge-Offline (OMO) platforms are widely used to enhance customer engagement, yet user participation often declines abruptly rather than gradually. Existing research largely assumes linear relationships between motivation, satisfaction, and continuance intention, overlooking the dynamic nature of user experience. Drawing on Self-Determination Theory (SDT) and Expectation Confirmation Theory (ECT), this study proposes a dynamic framework to explain how engagement evolves and when it collapses. Study 1 uses qualitative interviews to identify key mechanisms such as reward-driven motivation, expectation violation, and perceived unfairness. Study 2 introduces a multi-round experimental design simulating platform interactions under varying conditions. Study 3 proposes Bayesian Change Point Detection to capture non-linear shifts in psychological states. This research contributes to AI-driven e-commerce by revealing critical turning points in customer experience and offering insights for adaptive system design.

KEY WORDS

OMO, Gamification, Customer Experience, SDT, ECT, Change Point Detection, Dynamic Behavior

B1

Digital Transformation, Resource Allocation and Agri-Supply Chain Resilience.

Qianyi Wang, Ruyuan Huang (Shandong Technology and Business University)

This study addresses the resilience of the agricultural product supply chain under public emergencies, crucial for stable supply and food security. Using data from core enterprises (2014~2023), a resilience index covering learning, recovery, prediction, and resistance is constructed and measured via the entropy weight–TOPSIS method. Econometric analysis tests the effect of enterprise digital transformation. Key findings indicate: (1) digital transformation significantly enhances resilience; (2) effects are strongest in agro-processing and private sectors; (3) resource allocation mediates this impact. Recommendations target government, business, and industry to improve coordination, optimize operations, and safeguard food security.

KEY WORDS

Enterprise digital transformation, agricultural product supply chain resilience, resource allocation, mechanism effect

B1

Emotional Interaction and Conversion Value in Live Streaming E-commerce

Fen Qin (Taiyuan Univ. of Technology),
Yang Li (Shandong Normal Univ.) and Miao Feng (Shandong Management Univ.)

Emotional interaction in live streaming e-commerce involves multi-agent interactions, yet research on this topic remains scarce. Using 4,685 minutes of real-time data from 40 Taobao live sessions, this study applies sentiment analysis to capture emotions of streamers and viewers, and uses a PVAR model with impulse response analysis to examine dynamic relationships between streamer emotions, viewer emotions, assistant interventions, consumer engagement and purchase behavior. Findings reveal a one-way emotional contagion from streamers to consumers; viewer emotions can reversely influence streamers, while contagion among viewers is unsupported. Assistant interventions have a double-edged effect: boosting engagement but inhibiting sales when excessive. This study expands the application boundary of emotional contagion theory in live streaming e-commerce, improves the role theory system of the live streaming e-commerce ecosystem, and provides targeted guidance for live streaming operation practice.

KEY WORDS

Live Streaming E-commerce, Emotional Contagion, PVAR Model,
Live Streaming Assistant, Consumer Engagement

KITS·ICEC·KIISS
2026 통합 국제학술대회

B2 Corporate Strategy & Competency

Session Chair : Jinxin Yang (Hefei Tech. College)



B2

Resource Reconfiguration and Capability Regeneration: A Multilevel Mechanism Study on IT Innovation and Firm Performance

Chu Chao (Anhui University of Science and Technology), Yang Jinxin (Hefei Technology College)

In the era of deep integration of digitalization and intelligence, Information Technology (IT) has become the core driving force behind corporate innovation and performance enhancement. This paper, based on a systematic literature analysis method, conducts a comprehensive review of 46 empirical and theoretical studies published between 2015 and 2025, focusing on the relationship between organizational-level IT innovation and corporate performance. It reveals the theoretical evolution and paradigm shift in this field. The study finds that the Resource-Based View (RBV) and Dynamic Capabilities Theory (DCT) serve as the two core pillars for explaining the impact of IT innovation on performance. RBV emphasizes the scarcity and strategic value of resources, focusing on the performance contribution mechanism of IT as a heterogeneous resource. In contrast, DCT focuses on the process through which organizations perceive, capture, and reconfigure resources in uncertain environments, revealing how IT capabilities enable continuous innovation through dynamic adaptation. The integration of these two theories not only bridges the logical gap between "resources-capabilities" research, but also drives the shift in the research paradigm from static endowments to dynamic evolution. Further analysis reveals that the impact of IT innovation on corporate performance exhibits significant non-linearity and contextual dependency, with organizational agility, knowledge sharing, governance mechanisms, and environmental dynamics playing key mediating and moderating roles in this relationship. This paper theoretically constructs a chain-driven framework of "IT resources-IT capabilities-dynamic capabilities-innovation performance" and proposes a dual-mechanism model of resource reconfiguration and capability regeneration. On the practical level, the paper proposes a management action logic centered around "resource identification-capability building-governance alignment-performance realization". The research conclusions provide new theoretical foundations and practical insights for deepening the theoretical system of information technology innovation, optimizing digital investment decisions, and enhancing firms' sustainable competitive advantage.

KEY WORDS

Information Technology Innovation, Corporate Performance, Resource-Based View, Dynamic Capabilities Theory, Resource Reconfiguration, Capability Regeneration, Digital Transformation

Resource Reconfiguration and Capability Regeneration: A Multilevel Mechanism Study on IT Innovation and Firm Performance

Chu, C.a and Yang, J. X.b
a School of Economics and Management, Anhui University of Science and Technology
E-mail: 2024201606@aust.edu.cn
b School of Economics and Trade, Hefei Technology College
E-mail: jinxiny@aust.edu.cn

Abstract

In the context of deepening digital and intelligent integration, Information Technology (IT) has become a central driver of corporate innovation and performance. This paper systematically reviews 46 empirical and theoretical studies (2015–2025) on the relationship between organizational-level IT innovation and firm performance, tracing the theoretical evolution in this domain. The Resource-Based View (RBV) and Dynamic Capabilities Theory (DCT) emerge as the two foundational pillars. RBV emphasizes the strategic value of heterogeneous IT resources, while DCT illuminates how organizations sense, seize, and reconfigure resources under uncertainty, enabling sustained innovation through dynamic adaptation. Their integration bridges the "resources–capabilities" divide and shifts the paradigm from static endowments to dynamic evolution. The review further reveals that the IT innovation–performance link exhibits nonlinearity and contextual contingency, with organizational agility, knowledge sharing, governance mechanisms, and environmental turbulence serving as critical mediators and moderators. The paper advances a chain-driven framework of "IT resources→IT capabilities→dynamic capabilities→innovation performance" and a dual mechanism of resource reconfiguration and capability regeneration. Practically, it proposes a management logic of "resource identification → capability building→governance alignment→performance realization." These insights enrich the theoretical understanding of IT-enabled innovation and offer guidance for optimizing digital investment and sustaining competitive advantage.

Keywords: IT Innovation; Corporate Performance; RBV; DCT

Introduction

Corporate innovation performance generally refers to the outcomes an organization achieves during the innovation process, such as the launch of new products, the introduction of new processes, and the contribution of innovation to overall firm performance, with the overarching aim of enhancing competitiveness through organizational change. Although a substantial body of research has examined the role of information technology (IT) in innovation, theoretical integration in this domain remains relatively underdeveloped.

Advancing theoretical integration necessitates a systematic review of the existing literature. Current research exhibits several notable limitations. First, studies on information systems and organizational innovation have largely evolved in parallel, with a conspicuous lack of interdisciplinary synthesis. Second, distinct theoretical lenses—such as IT alignment, the resource-based view (RBV), and the capability perspective—remain fragmented and operate in relative isolation. Specifically, the RBV emphasizes the possession of IT resources, while capability research often conflates resource endowments with managerial actions. Meanwhile, the literature on IT–business alignment narrows the analytical scope, treating IT merely as a contextual backdrop and focusing predominantly on platform governance issues. As noted by previous scholars, research conducted prior to 2010 concentrated heavily on the internal processes of IT innovation development and implementation, paying scant attention to innovation initiation, commercialization, and the moderating influence of the external environment. Furthermore, the innovation mechanisms underlying emerging technologies—most notably artificial intelligence—have yet to be adequately explicated.

To address this gap, this paper undertakes a systematic screening of the extant literature to identify and delineate the key IT-related factors that shape organizational innovation. Additionally, by mapping the underexplored roles of emerging technologies and contextual non-linearities, the study offers a more nuanced understanding of IT-enabled innovation in the digital era.

Review of Related Concepts and Theories

The Resource-Based View

The RBV posits that competitive advantage derives from a firm's capacity to acquire and leverage resources that are valuable, rare, inimitable, and non-substitutable. Firm resources encompass assets, capabilities, organizational processes, information, and knowledge, and are broadly categorized within the RBV framework as either IT resources or non-IT organizational resources.

Within the information technology context, the resource-based view posits that IT enhances innovation performance by reinforcing the distinctiveness of organizational resources. Empirical evidence demonstrates that firms exhibiting superior IT capability not only attain higher profitability and lower cost structures relative to competitors—an outcome ascribed to elevated IT competence and consistently exceed industry-average profitability while sustaining comparatively reduced operating costs. Collectively, these findings substantiate that IT fortifies resource uniqueness and inimitability, thereby facilitating organizational innovation performance.

Dynamic Capabilities Theory

Dynamic capabilities theory builds upon the foundation of the RBV, emphasizing that firms cultivate competitive advantage through the dynamic orchestration and reconfiguration of unique resource portfolios. In contrast to the RBV's focus on static resource endowments, DCT adopts a process-oriented perspective, positing that capabilities enable firms to sense and seize emerging opportunities while realigning resources in response to market volatility. Dynamic capabilities are generally defined as "the capacity of an organization to purposefully create, extend, and modify its resource base."

In the context of information technology, DCT suggests that IT can augment a firm's dynamic capabilities—specifically in the dimensions of sensing, integrating, and innovating—thereby enhancing

innovation performance. Dynamic capabilities are conceptualized as encompassing five dimensions—sensing, learning, integrating, coordinating, and reconceptualizing—which collectively enable organizations to systematically recalibrate their operational routines and secure competitive advantage in turbulent markets.

Research Methods

Screening Process

This study adheres to the PRISMA guidelines for the systematic screening of the literature. Explicit inclusion and exclusion criteria were established, and multiple academic databases were queried to ensure both methodological rigor and replicability. Given the study's focus on the integration of primary theoretical and empirical research, review articles and editorials were excluded in the initial filtration phase. Subsequent hierarchical screening was conducted to retain only those studies situated at the organizational level of analysis; articles that did not address innovation or involved only a single entity were further eliminated. Figure 1 provides a schematic representation of the screening process.

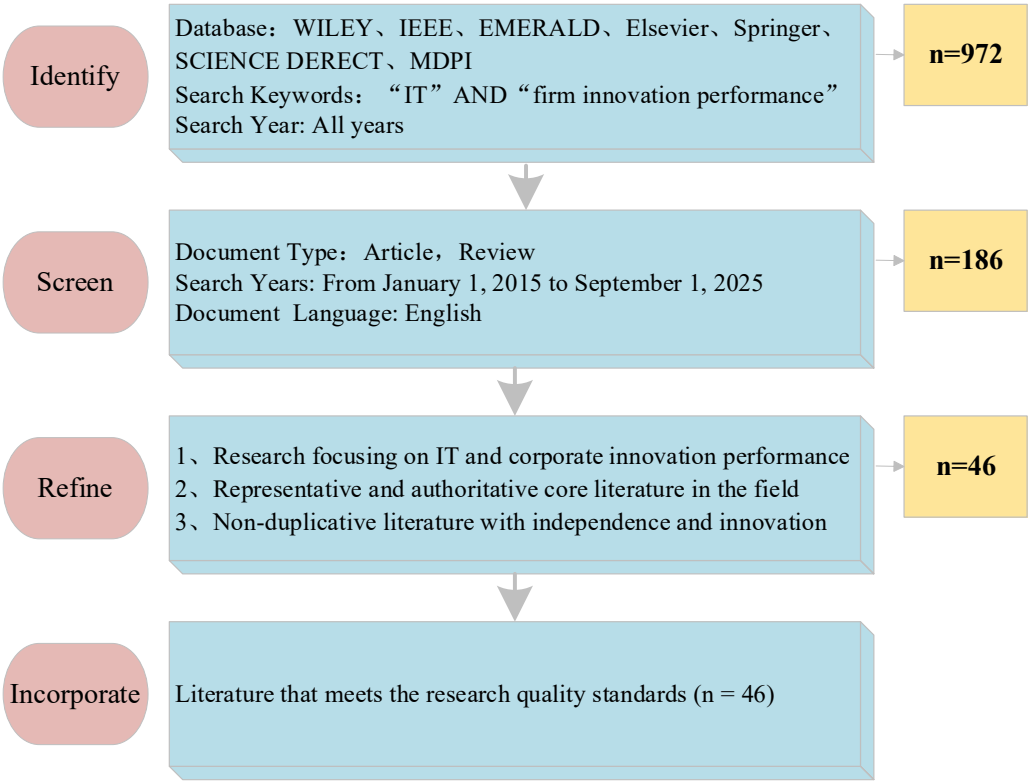


Figure 1：Screening Process

Literature Sources and Screening Strategies

The literature was sourced from Wiley, IEEE, ScienceDirect, Emerald, and Springer, supplemented by backward citation tracking. The search covered January 2015 to June 2025 using the string "IT innovation" AND "firm innovation performance." Articles were included if they examined this focal relationship, employed a relevant theoretical perspective, and appeared in peer-reviewed English-language journals. Studies lacking theoretical rigor or methodological quality were excluded.

After deduplication, abstract screening, and full-text assessment, 46 articles were retained from an initial 2,876 records.

Research and Analysis

Research Analysis Based on the RBV Theory

Extant research generally concurs that IT influences innovation through mediators such as absorptive capacity and knowledge sharing. Wai, Cheng-Hao, Liu, Nguyen, & Takeda (2020) identify IT-enabled support capabilities as mediating the strategic orientation–product innovation link; Al-Shami, Bakri, Adil, & Mamun (2021) confirm that IT capability fosters innovation via absorptive capacity; and J., A., & W. (2022) show that IT value addition enhances performance through technological innovation capability. Beyond linear mediation, studies reveal nonlinear dynamics: Wu, Zhang, & Tsai (2023) report an inverted U-shaped relationship between IT investment and collaborative innovation; C., M., S., & V. (2024) observe a U-shaped association between open innovation and performance; and Fichman & Melville (2014) caution that strategic and resource misalignments can sharply attenuate returns.

Another stream of literature examines the roles of IT governance, business–IT alignment, and usage intensity in shaping performance outcomes. Zhen, Xie, & Dong (2021) finds that top management support fosters IT ambidexterity through formal IT governance mechanisms, which in turn enhances organizational agility. Joshi, Benitez, Huygh, Ruiz, & De Haes (2022) confirms that IT governance process capabilities influence business performance indirectly via their impact on IT performance. Ilmudeen, Bao, & Zhang (2022) demonstrates that business–IT alignment partially mediates the relationship between IT investment management and firm performance. Furthermore, Bourdeau observes that both IT usage intensity and innovation culture exert a positive influence on innovation intensity, thereby elevating overall performance.

In recent years, the intersection of IT and green innovation has emerged as a prominent research frontier. Zhou finds that global diversification significantly promotes green innovation; Han confirms that digital transformation exerts an inverted U-shaped influence on green innovation; and Du demonstrates that IT application can directly drive green innovation outcomes. Wu (2017) and Do, Do, & Nguyen (2024) provide supplementary evidence from a sustainability perspective. Eom, Zaza, & Erskine (2025) reports that IT project leadership positively affects project performance. Additionally, Dani, Gandhi, & Sharma (2023) employs a grounded theory approach to identify eight drivers of internal IT innovation, offering a valuable qualitative complement to the predominantly quantitative literature.

Research Analysis Based on DCT

With respect to IT-enabled collaborative innovation and supply chain digitalization, Wang, Zhao, Chi, & Li (2017) finds that digital platform capability exerts a fully mediated effect on competitive performance through collaborative innovation capability, with relational governance positively moderating and formal governance negatively moderating this relationship. Ayala et al. (2020) identifies knowledge management systems and collaborative work systems as critical enablers of collaborative new product development. Lyu, Guo, & Lin (2022) demonstrates that information systems integration partially mediates the influence on supply chain ambidexterity via dynamic capabilities. Zhao, Hong, & Lau (2023) reports that supply chain digitalization partially mediates

performance outcomes through enhanced responsiveness and recovery capabilities. Furthermore, Lin proposes a fit mechanism between digital capabilities and governance structures that positively influences supply chain resilience.

Regarding IT capability and digital transformation, W. & S. (2025) finds that IT knowledge, IT infrastructure, and IT operational capabilities all positively influence organizational agility, with organizational learning serving as a moderating factor.Pang, Zhang, Chen, Chen, & Yang (2025) confirms that IT intellectual capital positively affects digital transformation, wherein social capital acts as a mediator, while market turbulence exerts differential moderating effects depending on its direction. R. & A. (2022) reports that entrepreneurial orientation and technological orientation positively influence organizational performance through the mediating mechanism of IT demand management.

In the domain of management innovation and green innovation, E. & R. (2023) finds that management innovation negatively moderates the relationship between technological innovation and firm performance. Furthermore, research confirms that green management innovativeness positively influences green innovation performance through the mediating mechanism of green ambidexterity. Liu, Ning, & Gao (2024) demonstrates that digital ecosystem embeddedness enhances innovation performance via ambidextrous capabilities, with strategic flexibility exerting a positive moderating effect.

Research Analysis Based on RBV and DCT

With respect to IT capability, digital leadership, and organizational agility, Tsou & Cheng (2018) finds that IT capability and organizational learning fully mediate service innovation performance through organizational agility. Ravichandran (2018) confirms that IS capabilities enhance agility and firm performance via digital platforms. Weng, Wang, De Lurgio II, & Schuetz (2024) reports that IT capability fully mediates financial performance through value proposition innovation. Hu & Wang (2021) shows that IT usage breadth and frequency positively affect innovation performance, though manufacturing flexibility mediates only frequency. Benitez, Arenas, Castillo, & Esteves (2022) demonstrates that digital leadership influences innovation performance through platform digital capabilities. Chen, Lu, & Tongnoy (2025) propose a hierarchical framework that traces the evolution of IT capabilities from asset-level resources to dynamic capabilities. Furthermore, research indicates that strategic orientation positively influences both performance and innovation culture, with market turbulence strengthening these relationships.

In green innovation and resilience, Adomako & Nguyen (2023) finds green creativity affects green product performance through responsible innovation. Abbas confirms IT capability influences green innovation via green supply chain integration, with customer integration mediating strongest. Aftab, Stan, Srivastava, Wei, & Abid (2025) shows digital leadership affects performance through big data analytics and green innovation, moderated by AI change readiness. Sun, Mou, Zhu, & Jeyaraj (2025)demonstrate that digital technology empowerment enhances resilience through ambidextrous innovation, that digitalization influences innovation value appropriation via supply chain resilience, and that firms with IT backgrounds are better positioned to advance green innovation.

Regarding external environment and non-market strategy, Li, Gao, & Kwok (2024) finds that non-market strategies enhance innovation performance, positively moderated by internal capabilities and negatively moderated by the external IT environment. Trieu, Nguyen, & Vrontis (2025) reports that IT adoption indirectly affects performance through innovation and resilience, with government

support playing a positive role. Qi et al. (2023) finds that the synergy between M&A and internal innovation influences product differentiation. These findings predominantly derive from quantitative studies, with a limited number incorporating interviews or mixed methods.

Findings

In terms of theoretical fit, RBV prevails due to the VRIO framework's operational utility in assessing IT resource value, whereas DCT better explains IT-driven dynamic adaptation. Studies integrating both perspectives reveal their complementarity, indicating that neither lens alone fully captures IT innovation mechanisms. rather, an integrated approach provides more comprehensive coverage of the logic chain spanning "resource foundation → dynamic transformation → performance manifestation."

Recommendations

This study offers several actionable insights for managers. First, firms should cultivate dynamic capabilities—sensing, learning, integrating, and reconfiguring—to adapt IT to shifting market conditions, rather than treating IT as a static asset. Second, aligning IT governance with strategic intent is critical, as misalignment can erode innovation returns. Third, the observed non-linearities caution against indiscriminate IT investment, highlighting the need to calibrate spending with organizational absorptive capacity and environmental volatility.

Conclusion

Despite the considerable scholarly attention devoted to the impact of information technology on organizational innovation performance, notable gaps persist in the extant literature. Although the RBV and DCT have been widely applied, their inherent limitations constrain the explanatory breadth of current theoretical accounts. Moreover, the role of emerging technologies—most notably artificial intelligence and blockchain—remains underexplored. Future research should therefore broaden the theoretical repertoire by integrating diverse conceptual frameworks, delve more deeply into the innovation implications of nascent technologies, and examine the synergistic interplay between resources and capabilities. Such endeavors will yield a more comprehensive understanding of the mechanisms through which information technology propels organizational innovation performance.

References

Adomako, S., & Nguyen, N. P. (2023). Green creativity, responsible innovation, and product innovation performance: A study of entrepreneurial firms in an emerging economy. *Business Strategy and the Environment*, 32(7), 4413-4425. doi: https://doi.org/10.1002/bse.3373

Aftab, J., Stan, M. R., Srivastava, M., Wei, F., & Abid, N. (2025). The Impact of Digital Leadership on Performance: Examining the Roles of Big Data Analytical Capabilities, Green Innovation, and AI Change Readiness in Italian SMEs. *Business Strategy and the Environment*, 34(6), 6864-6884.

Al-Shami, S., Bakri, M. H., Adil, H., & Mamun, A. A. (2021). Information technology competencies as antecedents for absorptive capacity and innovation capabilities in a high-tech industry. *Foresight*, 24(5), 565-585. doi: 10.1108/FS-12-2020-0129

Ayala, N. F., Dain, M. L., Merminod, V., Gzara, L., Enrique, D. V.,... Frank, A. G. (2020). The contribution of IT-leveraging capability for collaborative product development with suppliers. *J. Strateg. Inf. Syst.*, 29, 101633

Benitez, J., Arenas, A., Castillo, A., & Esteves, J. (2022). Impact of digital leadership capability on innovation performance: The role of platform digitization capability. *Information & Management*, 59(2), 103590. doi: https://doi.org/10.1016/j.im.2022.103590

C., T. L., M., B., S., M., & V., D. (2024). Open Innovation Where it Really Matters: The U-Shaped Relationship Between Relative Open Innovation and Innovation Performance in Developing Countries. *IEEE Transactions on Engineering Management*, 71, 15540-15554.

Chen, D., Lu, K., & Tongnoy, S. (2025). The Role of IT Capability and Knowledge Sharing in Open Innovation Adoption of an Organization. *Information Systems Management*, 42(3), 330-345.

Dani, M. V., Gandhi, A. V., & Sharma, A. (2023). Understanding drivers of innovation in information technology companies using a grounded theory approach. *International Journal of Innovation Science*, 15(5), 817-838. doi: 10.1108/IJIS-07-2022-0119

Do, H., Do, N. B., & Nguyen, T. M. (2024). Unveiling the role of sustainability practices, co-innovation and entrepreneurial orientation on SMEs performance: An empirical study in developing country. *Business Strategy & Development*, 7(4), e70022.

E., A. H., & R., A. C. M. (2023). Management Innovation in an Emerging Economy: An Analysis of Its Moderating Effect on the Technological Innovation–Performance Relationship. *IEEE Transactions on Engineering Management*, 70(1), 128-141. doi: 10.1109/TEM.2021.3052746

Eom, M. T., Zaza, S., & Erskine, M. A. (2025). Overcoming IT Project Challenges: An Integrated Model of Characteristics, Leadership, and Complexity. *Information Systems Management*, 42(3), 265-289. doi: 10.1080/10580530.2024.2387582

Fichman, R. G., & Melville, N. P. (2014). How Posture-Profile Misalignment in IT Innovation Diminishes Returns: Conceptual Development and Empirical Demonstration. *Journal of Management Information Systems*, 31(1), 203-240. doi: 10.2753/MIS0742-1222310109

Hu, S., & Wang, X. (2021). IT Usage and Innovation Performance of SMEs in China: A New Perspective. *Discrete Dynamics in Nature and Society*, 2021(1), 5512933.

Ilmudeen, A., Bao, Y., & Zhang, P. (2022). Investigating the Mediating Effect of Business-IT Alignment between Management of IT Investment and Firm Performance. *Information Systems Management*, 40, 208-228

J., Y., A., T., & W., C. (2022, 2022-1-1). *An empirical study on the impact of technological innovation on enterprise performance based on IT value-added theory*. Paper presented at the 2022 European Conference on Natural Language Processing and Information Retrieval (ECNLPIR).

Joshi, A., Benitez, J., Huygh, T., Ruiz, L., & De Haes, S. (2022). Impact of IT governance process capability on business performance: Theory and empirical evidence. *Decision Support Systems*, 153, 113668. doi: https://doi.org/10.1016/j.dss.2021.113668

Li, Y., Gao, S., & Kwok, R. C. (2024). Nonmarket strategy and innovation performance in pharmaceutical industry: the moderating effect of internal capability and external IT environment. *Chinese Management Studies*, 18(5), 1492-1513. doi: 10.1108/CMS-02-2023-0065

Liu, J., Ning, L., & Gao, Q. (2024). Research on the mechanism of digital innovation ecosystem embeddedness on the digital innovation performance of complementary enterprises: evidence from China. *Kybernetes*. doi: 10.1108/K-12-2023-2709

Lyu, T., Guo, Y., & Lin, H. (2022). Understanding green supply chain information integration on supply chain process ambidexterity: The mediator of dynamic ability and the moderator of leaders' networking ability. *Frontiers in Psychology*, 13. doi: 10.3389/fpsyg.2022.1088077

Pang, B., Zhang, M., Chen, Y., Chen, M., & Yang, R. (2025). Unlocking digital transformation: the

impact of IT intellectual capital and market turbulence. *Internet Research*.
Qi, K., Han, S., Jeong, Y., Rhamati, P., Susarla, A.,... Sambamurthy, V. (2023). Product Differentiation in IT Industries: The Interplay between M&As and In-House Innovation and Exploitative Innovation Orientation. *Journal of the Association for Information Systems*, 24, 1102-1131. doi: 10.17705/1jais.00828
R., W., & A., L. (2022). Strategic IT Demand Management for Business and Innovation Organization. *Journal of Mobile Multimedia*, 18(6), 1851-1878. doi: 10.13052/jmm1550-4646.18617
Ravichandran, T. (2018). Exploring the relationships between IT competence, innovation capacity and organizational agility. *The Journal of Strategic Information Systems*, 27(1), 22-42.
Sun, Y., Mou, S., Zhu, M., & Jeyaraj, A. (2025). Impact of digital technology affordance on organizational resilience: An ambidextrous innovation perspective using a mixed-methods approach. *Journal of Business Research*, 197, 115450. doi: https://doi.org/10.1016/j.jbusres.2025.115450
Trieu, H. D. X., Nguyen, P. V., & Vrontis, D. (2025). The interplay of IT adoption, government support and firm performance: mediating roles of innovation and resilience in Vietnamese SMEs. *Journal of Asia Business Studies*, 19(4), 935-953. doi: https://doi.org/10.1108/JABS-12-2024-0690
Tsou, H., & Cheng, C. C. J. (2018). How to enhance IT B2B service innovation? An integrated view of organizational mechanisms. *Journal of Business & Industrial Marketing*, 33(7), 984-1000. doi: 10.1108/JBIM-07-2017-0175
W., Y., & S., W. K. (2025). The Impact of IT Capabilities on Organizational Agility With the Moderating Effect of Organizational Learning. *IEEE Access*, 13, 30915-30935. doi: 10.1109/ACCESS.2025.3542575
Wai, W. J. K., Cheng-Hao, S. C., Liu, G., Nguyen, B., & Takeda, S. (2020). IT-based product innovation strategies for small firms. *Information Technology & People*, 33(6), 1489-1514. doi: https://doi.org/10.1108/ITP-07-2018-0343
Wang, F., Zhao, J., Chi, M., & Li, Y. (2017). Collaborative innovation capability in IT-enabled inter-firm collaboration. *Industrial Management & Data Systems*, 117(10), 2364-2380. doi: 10.1108/IMDS-09-2016-0392
Weng, Q., Wang, D., De Lurgio II, S., & Schuetz, S. (2024). How do small-to-medium-sized e-commerce businesses stay competitive? Evidence on the critical roles of IT capability, innovation and multihoming. *Internet Research*, 35(1), 126-151. doi: 10.1108/INTR-01-2023-0061
Wu, G. (2017). Effects of Socially Responsible Supplier Development and Sustainability-Oriented Innovation on Sustainable Development: Empirical Evidence from SMEs. *Corporate Social Responsibility and Environmental Management*, 24(6), 661-675. doi: https://doi.org/10.1002/csr.1435
Wu, W., Zhang, P., & Tsai, S. (2023). IT investment and corporate collaborative innovation: the moderating role of the top management team's educational background and absorptive capacity. *Internet Research*, 34(3), 1017-1045. doi: 10.1108/INTR-12-2021-0899
Zhao, N., Hong, J., & Lau, K. H. (2023). Impact of supply chain digitalization on supply chain resilience and performance: A multi-mediation model. *International Journal of Production Economics*, 259, 108817. doi: https://doi.org/10.1016/j.ijpe.2023.108817
Zhen, J., Xie, Z., & Dong, K. (2021). Impact of IT governance mechanisms on organizational agility and the role of top management support and IT ambidexterity. *International Journal of Accounting Information Systems*, 40, 100501. doi: https://doi.org/10.1016/j.accinf.2021.100501

B2

Pragmatism over Greenery: Why Utilitarian Benefits Drive Smart Locker Adoption in Vietnam's Last-Mile Delivery.

Minh-Trang Hoang, Chih-Hsiung Hu
(Department of Industrial Management, National Formosa University, Taiwan),
Tien-Dat Nguyen (Department of Smart Industry Technology Research and Development, National Formosa University, Taiwan & Faculty of Economics and Management, Van Hien University, Vietnam)
and Chih-Chin Liang (Department of Business Administration, National Formosa University, Taiwan).

This study examines the adoption of self-collection services (SCS), specifically smart parcel lockers, in Vietnam's last-mile delivery sector. Integrating the Theory of Planned Behavior and Technology Readiness Index, we analyze data from 372 experienced consumers using PLS-SEM. The findings reveal that utilitarian attributes-convenience, reliability, and privacy-are the primary drivers of perceived behavioral control and positive attitudes. Paradoxically, environmental concern failed to significantly influence adoption intention, highlighting a pronounced "Value-Action Gap"in this emerging market. These results suggest that for Vietnamese consumers, pragmatic personal benefits heavily override ecological consciousness. The research provides actionable insights for logistics providers to pivot marketing strategies toward 24/7 autonomous retrieval and service efficiency rather than abstract green messaging.

KEY WORDS

Self-collection services, Smart parcel lockers, Last-mile delivery, Theory of planned behavior

B2

An Investigation into the Mechanisms of Enterprise Digital Transformation Through the Lens of New Institutional Theory.

Yingzhao He, Chengchi Zhou (Northwest Normal University)

Based on New Institutional Theory and Institutional Entrepreneurship Theory, this study constructed a theoretical model in which the external institutional pressure of an enterprise affects the internal initiative factors, and then the state of internal institutional change of the enterprise, ultimately influencing the process of digital transformation of the enterprise. The model was tested by PLS-SEM. Research has found that institutional pressure has a positive effect on the digital transformation of enterprises, and this effect is not a direct one but is achieved through the mediating role of digital strategy, digital leadership and digital readiness.

KEY WORDS

Digital transformation, Institutional pressure, Digital readiness, Digital strategy, Digital leadership

B2

Interpretable Machine Learning for Predicting Employee Attrition in SMEs

Fotima Yusupova, So Ra Park (College of Business Administration, Chonnam National University).

In condition of Small and Medium sized Enterprises, system strongly relies on each employee, which loosing even one of them can cause strong affect on daily productivity. In this study we compared several machine learning models specifically glass-box and black-box, for prediction attrition and questioned how how understandable their reasonings on predicting. The analysis done by using a well-known employee dataset and creating an SME-like setting to reflect more limited data conditions. For this study we followed a careful evaluation process to avoid data leakage and to ensure fair comparison across models.

The findings show that the Explainable Boosting Machine performs very closely to the strongest black-box model while offering clearer and more direct interpretation. The explanation analysis also shows partial agreement between post-hoc and built-in explanations, while the glass-box model reveals additional interaction patterns. Overall, the results suggest that Explainable Boosting Machine offers a useful balance between predictive strength and transparency for, managing the internal customer, decision-making in SMEs.

KEY WORDS

Employee attrition prediction, Explainable AI, Interpretable machine learning, SME decision support, SHAP, Explainable Boosting Machine, Internal customer

Interpretable Machine Learning for Predicting Employee Attrition in SMEs

Yusupova Fotima Javlon kizi[a] and Park, So Ra[b]
[a] Department of Digital Service Convergence, Chonnam National University
E-mail: yusupova_fotima@jnu.ac.kr
[b] College of Business Administration, Chonnam National University
E-mail: doorae0207@jnu.ac.kr

Abstract

In condition of Small and Medium sized Enterprises (SMEs), system strongly relies on each employee, which loosing even one of them can cause a strong effect on daily productivity. In this study we compared several machine learning models both glass-box and black-box models, for predicting attrition and questioned how understandable their reasonings on predicting. The analysis done by using a well-known employee dataset and creating an SME-like setting to reflect more limited data conditions. For this study we followed a careful evaluation process to avoid data leakage and to ensure fair comparison across models. The findings show that among glass-box models, the Explainable Boosting Machine's performance match very closely to the strongest black-box model, while offering clearer and more direct interpretation. The explanation analysis also shows partial agreement between post-hoc and built-in explanations, while the glass-box model reveals additional interaction patterns. Overall, the results suggest that Explainable Boosting Machine offers a useful balance between predictive strength and transparency for decision-making, managing the internal customer, in SMEs.

Keywords: Employee attrition prediction; Explainable AI; Interpretable machine learning; SME decision support; SHAP; Explainable Boosting Machine, Internal customer

Introduction

Employee attrition can be especially damaging for small and medium-sized enterprises (SMEs), where even the loss of one employee can place significant pressure on a small workforce (Baharim et al. 2024; Fernet et al. 2016). Despite this, many SMEs have limited analytical tools, HR systems, and data infrastructure to identify attrition risk at an early stage (Alsaadi et al. 2022; Baharim et al. 2024; Fernet et al. 2016). Burnout is closely linked to job demands, strain, and withdrawal processes, and it often emerges before employees' decision to leave (Bakker et al. 2008; Deary et al. 2003; Demerouti et al. 2001). However, burnout and attrition are often studied separately (Alqahtani et al. 2024; Heno Merlin C. P. S. 2025; Kumar et al. 2023; Talebi et al. 2025).

Most existing machine learning research on attrition relies on black-box models, with interpretability added post-hoc utilizing methods such as SHAP or LIME (Al-Ali et al. 2026; Alqahtani et al. 2024; Narkbunnum and Hinthaw 2025; Pavithran and Vadivel 2026; Talebi et al. 2025). Many of these studies used the IBM HR dataset to compare different models, and they do not reflect on the practical limitations faced by SMEs, e.g. small avaialbe data-sets (Alqahtani et al. 2024; Baharim et al. 2024; Kumar et al. 2023; Talebi et al. 2025). Therefore, direct comparisons between inherently interpretable models and black-box models with post-hoc explainability under small-sample, and leakage-free conditions remain limited (Alqahtani et al. 2024; Guerranti and Dimitri 2022; Narkbunnum and Hinthaw 2025; Talebi et al. 2025).

This study addresses the relationship between burnout and attrition in SMEs in three ways. First, it benchmarks five model families, including the glass-box Explainable Boosting Machine (EBM), against black-box alternatives under both full-data and SME-simulated small-sample conditions using a leakage-free evaluation protocol. Second, it compares intrinsic and post-hoc interpretability by examining feature importance, effect direction, and explanation stability across runs. Third, it shows that EBM can achieve prediction performance close to strong black-box models while offering clearer explanations and useful interaction effects, making it a practical choice for SME decision-makers.

Related work

Employee attrition and burnout have been explained through several complementary theoretical lenses. Classical turnover-process views, often associated with Mobley, describe leaving as a gradual process in which dissatisfaction develops into withdrawal, job search, and eventual exit. In parallel, the Job Demands-Resources (JD-R) model explains burnout as the result of imbalance between job demands and available resources, resulting exhaustion, disengagement, and turnover intention more likely (Demerouti et al. 2001; Bakker et al. 2008; Scholze and Hecker 2024). Burnout research further shows that psychological strain and exhaustion's role onweakening employee attachment and work functioning over time (Fernet et al. 2016). Retention-oriented perspectives such as job embeddedness add an important counterpoint by focusing on why they stay instead of why they leave. by enhancing fit, and sacrifice within the organization; this broader retention logic also appears in later review research on turnover (Kumar et al. 2023; Alqahtani et al. 2024; Talebi et al. 2025).

Recent research has increasingly incorporated machine learning to employee attrition prediction, and utilizing the IBM HR Analytics Employee Attrition and Performance dataset has become a norm in this stream. Comparative studies have repeatedly tested Logistic Regression, Decision Trees, Random Forest, SVM, XGBoost, and deep learning models on the IBM-based data, often reporting strong predictive performance while relying on the same 1,470-observation dataset (Al-Ali et al. 2026; Guerranti and Dimitri 2022; Kumar et al. 2023; Raza et al. 2022). This pattern has helped standardize the model evaluation, but it has also narrowed the empirical base of the field. Overall, recent reviews suggest that ensemble and boosting methods often outperform simpler models on headline accuracy, while simpler models remain attractive due to their interpretability and their practicability in the deployment in HR settings (Kannimuthu and Premalatha 2024; Pavithran and Vadivel 2026).

As predictive performance improved, interpretability became a central issue in HR analytics because managers need not only a risk score but also understandable reasons behind it. Much of the recent literature therefore relies on post-hoc explainability methods, especially SHAP and broader LIME-style logic, to identify influential features and to communicate global or individual-level drivers of attrition (Al-Ali et al. 2026; Pavithran and Vadivel 2026). At the same time, the literature is gradually moving toward adopting glass-box models, intrinsically interpretable approaches, revealing nonlinear effects and interaction patterns more directly than post-hoc attribution alone. However, direct comparisons between post-hoc and glass-box interpretability remain limited in attrition studies (Manafi Varkiani et al. 2025; Narkbunnum and Hinthaw 2025; Talebi et al. 2025).

While most existing models are evaluated under relatively data-rich conditions. small SME HR datasets are available limiting overall analytical capacity (Alqahtani et al. 2024; Kumar et al. 2023; Talebi et al. 2025). Yet most explainable attrition studies still benchmark models on IBM or other larger datasets rather than testing whether performance and explanations remain reliable when data become scarce. To address the gap this research compares glass-box and post-hoc explainability under both standard IBM conditions and an SME-simulated low-data setting, with attention to minority-class performance, stability, and explanation quality.

Methodology

This study aims to compare both the predictive performance and interpretability of different machine learning models for employee attrition prediction under two settings: a full-data setting and an SME-like setting (Govindarajan et al. 2025; Guerranti and Dimitri 2022; Manafi Varkiani et al. 2025;

Narkbunnum and Hinthaw 2025). The methodology followed five main stages: data preparation, preprocessing and encoding, experimental design and validation, model development and evaluation, and explainability and stability analysis. All preprocessing and modeling steps were applied in a consistent way before carrying out the interpretation stage. A summary of the dataset and experimental settings is shown in Figure 1.

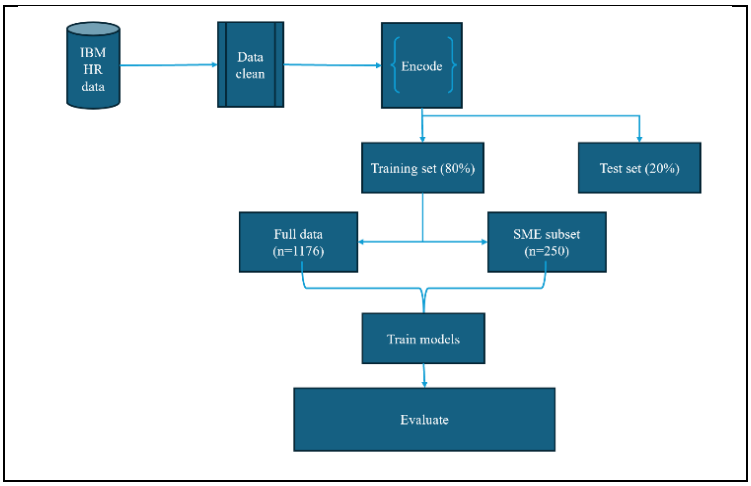


Figure 1. Research Flow

3.1 Data Description

This study used the IBM HR Analytics Employee Attrition dataset, which contains 1,470 employee records, 35 original variables, and an attrition rate of 16.1% (Alqahtani et al. 2024; Kumar et al. 2023; Manafi Varkiani et al. 2025; Talebi et al. 2025). The dataset was selected because it provides structured employee-level information related to attrition, including demographic, job-related, salary, and workplace factors. This makes it suitable for comparing both machine learning performance and interpretability methods within a single setting (Govindarajan et al. 2025; Manafi Varkiani et al. 2025; Mansor et al. 2021; Raza et al. 2022).

To simulate SME-like conditions, 250 records were randomly drawn from the training data, producing a smaller dataset with an attrition rate of 18.4%. The held-out test set contained 294 observations with an attrition rate of 16.0%. The same test set was used in both the full-data and SME-like settings so that any difference in results could be attributed to training data size rather than changes in the evaluation sample. This design allowed the dataset to serve both as a benchmark and as a simulation of limited-data conditions often faced by SMEs.

3.2 Data Preprocessing and Feature Construction

Before model training, the dataset was preprocessed to remove non-informative variables and convert the data into a format suitable for machine learning, following common practices in attrition prediction studies (Govindarajan et al. 2025; Manafi Varkiani et al. 2025; Mansor et al. 2021; Raza et al. 2022). Four columns were removed: EmployeeCount, Over18, StandardHours, and EmployeeNumber, as they had no meaningful relationship with attrition.

Next, the target variable Attrition was converted into binary form for classification (Attrition = 1, No Attrition = 0). Categorical variables were then transformed using one-hot encoding, while all dummy variables were retained. This preserved category-level information and made the data compatible with both linear and tree-based models. After encoding, the final dataset contained 51 predictor variables.

The aim of preprocessing was to standardize the dataset while maintaining interpretability. Rather than reducing the feature set too aggressively, the study retained a broad range of HR-related variables so that the models could be compared in terms of how they identified key attrition drivers.

3.3 Experimental Design

The dataset was divided using a stratified 80/20 train-test split in order to preserve the class distribution in both training and test sets. Because attrition represented only a small share of the observations, stratification was necessary to ensure balanced evaluation. The test set was kept fixed throughout the study and used only for final evaluation, while all resampling and validation steps were performed on the training portion only.

To simulate the limited-data environment common in SMEs, a smaller training subset of 250 observations was randomly selected from the original training data. Using the same held-out test set for both the full-data and SME-like conditions allowed a fair comparison between the two training settings.

Because the attrition class was imbalanced, SMOTE was used during model development. However, to prevent data leakage, it was applied only within the training folds of cross-validation and never before the train-test split or on validation and test data. Overall, the experimental design emphasized fairness, reproducibility, and a clear separation between training, validation, and testing.

3.4 Model Development and Evaluation Metrics

Five machine learning models were selected for the study: Logistic Regression (LR), Decision Tree (DT), Explainable Boosting Machine (EBM), XGBoost, and Random Forest (RF). These models were chosen to represent both simpler interpretable methods and more flexible ensemble approaches, allowing comparison of predictive performance across different levels of interpretability (Alqahtani et al. 2024; Govindarajan et al. 2025; Guerranti and Dimitri 2022; Kumar et al. 2023; Talebi et al. 2025).

EBM served as the main glass-box model. Similar to a generalized additive model, it can also detect pairwise interactions and directly interpret feature effects. In contrast, XGBoost and Random Forest were treated as stronger but less transparent models and were later interpreted using post-hoc explainability tools (Al-Ali et al. 2026; Manafi Varkiani et al. 2025; Narkbunnum and Hinthaw 2025; Pavithran and Vadivel 2026).

Model training was conducted using 5-fold stratified cross-validation on the training data, with SMOTE applied within each training fold. Performance was evaluated using minority-class F1-score, PR-AUC, ROC-AUC, and Brier score, with primary emphasis on PR-AUC and minority-class F1-score because of class imbalance. A fixed random seed of 42 was used for all models to support reproducibility.

3.5 Explainability and Stability Analysis

After model development, a separate explainability analysis was conducted to compare post-hoc and intrinsic interpretability. For XGBoost and Random Forest, SHAP TreeExplainer was used to identify important features and show whether they increased or reduced attrition risk, helping reveal the main drivers of predictions (Al-Ali et al. 2026; Manafi Varkiani et al. 2025; Narkbunnum and Hinthaw 2025; Pavithran and Vadivel 2026; Talebi et al. 2025). For EBM, interpretation came directly from the model's own outputs, including feature importance, term scores, and shape functions, allowing examination of effect direction and nonlinear patterns. To assess reliability, training and feature ranking were repeated with multiple random seeds, and rank variation was measured (Guerranti and Dimitri 2022; Manafi Varkiani et al. 2025). EBM main effects were also separated from interactions. All outputs were saved for transparent, reproducible, and clear reporting.

Results and Discussion

4.1 Predictive performance

Under the full-data condition, XGBoost achieved the best predictive performance, with the highest PR-AUC (0.5299) and minority-class F1 score (0.4872). EBM ranked second and remained very close, reaching a PR-AUC of 0.5247 and an F1 score of 0.4810. This small gap shows that the glass-box model performed nearly as well as the strongest black-box model. In contrast, Decision Tree produced the

lowest PR-AUC (0.3516), indicating weaker ranking ability under class imbalance. Under the SME-simulated condition (n = 250), PR-AUC declined across all models, showing reduced discrimination with limited training data. XGBoost remained the top model (0.4672), while EBM stayed close (0.4516). These findings suggest that EBM retained competitive performance even in the smaller-data setting.

Table 1

Model	Type	F1 (Full)	PR_AUC (Full)	F1 (SME)	PR_AUC (SME)
Logistic Regression	Base line	0.4286	0.4504	0.4062	0.3961
Decision Tree	Transparent	0.4423	0.3516	0.3486	0.2265
EBM	Glass-Box	0.481	0.5247	0.4	0.4516
XGBoost	Black-Box	0.4872	0.5299	0.4176	0.4672
Random Forest	Black-Box	0.3529	0.5091	0.3704	0.4314

The SHAP summary plot shows that XGBoost is driven mainly by work conditions and employee experience. *OverTime_No* is the strongest factor: employees who work overtime are more likely to leave, while those who do not are less likely to leave. *JobSatisfaction* and *EnvironmentSatisfaction* show the same pattern, with lower satisfaction increasing attrition risk. Shorter tenure, especially in the current role or with the current manager, is also linked to higher risk. Other variables have weaker effects. *NumCompaniesWorked* suggests that frequent past job changes increase attrition risk, while higher *MonthlyIncome* and *JobLevel* generally reduce it. Some categorical variables show mixed effects without a clear direction.

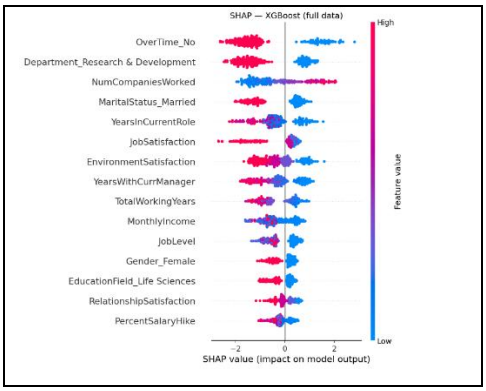


Figure 2. SHAP XGBoost Full data

4.2 Explainability quality

The results show that XGBoost and EBM perform almost identically on the full dataset, with only very small differences in minority-class F1 (gap ≈ 0.0062) and PR-AUC (gap ≈ 0.0052). This suggests that EBM can match the predictive quality of a strong black-box model while remaining interpretable. Under the SME setting, performance declines across all models, showing sensitivity to limited data. Decision Tree degrades the most (PR-AUC drop ≈ 0.125), indicating weak stability in small samples. In contrast, XGBoost and EBM remain relatively robust, with XGBoost only slightly outperforming EBM in PR-AUC under SME conditions (0.4672 vs 0.4516).

Table 2

Rank	SHAP_XGB	EBM
Top-5 features	*OverTime_No, Department_R&D, NumCompaniesWorked,*	*OverTime_No, YearsWithCurrManager, MaritalStatus_Married,*

	MaritalStatus_Married, YearsInCurrentRole	*JobSatisfaction, YearsInCurrentRole*
Mean rank variance	4.624	6.432
Overlap	3/5	3/5
Interaction terms	0	8
OverTime_No direction	- (protective)	- (protective, bin 2 = -0.639)

From an interpretability perspective, the models show moderate agreement, with 3 of the top 5 features overlapping. XGBoost explanations are more stable (lower variance ≈ 4.624), showing lower rank variance than EBM(≈ 6.432). However, EBM offers added value through transparent interaction effects, with the strongest found between *HourlyRate* and *EducationField_Life Sciences*. The negative score for *OverTime_No* (-0.639) also confirms a consistent protective effect against attrition. Overall, while XGBoost provides slightly greater stability, EBM combines similar predictive performance with richer and more interpretable insights, making it a strong option for SME decision support.

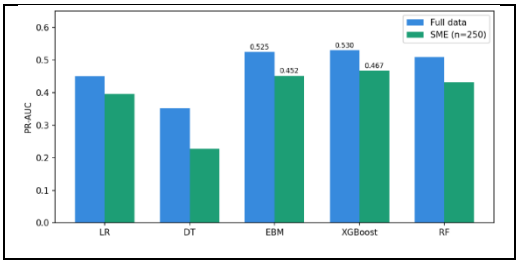


Figure 3. PR_AUC Comparison

The PR-AUC results show clear differences between full-data and SME conditions. With the full dataset, XGBoost (0.530) and EBM (0.525) perform best, showing strong ability to detect attrition despite class imbalance. Random Forest also performs well (0.510), while Logistic Regression is moderate (0.450). Decision Tree performs worst (0.352), suggesting weaker generalization.

When training is reduced to the SME subset (n = 250), all models decline. However, XGBoost (0.467) and EBM (0.452) remain the most stable, indicating better robustness with limited data. Logistic Regression drops slightly, while Decision Tree falls sharply (0.226), confirming poor stability. Random Forest also declines noticeably. Overall, EBM offers the best balance between predictive performance and interpretability, while XGBoost achieves the highest accuracy but depends on post-hoc explanations.

4.3 Discussion

The results show that EBM performs very closely to XGBoost under full-data conditions, with only small differences in performance. This suggests that EBM is competitive with a strong black-box model. Under the SME-simulated setting, EBM achieves the highest PR-AUC, showing that it remains reliable even when data is limited. This is important because many SMEs do not have large HR datasets, so models that perform well under small-sample conditions are more practical.

The explanation results show partial agreement across methods. The overlap of 3 out of 5 important features, together with the consistent directional effect of OverTime_No, supports the credibility of EBM as a transparent alternative to black-box models. At the same time, disagreement on two features should not be treated only as a weakness. It likely reflects differences in how each method assigns importance and captures variance. For HR practitioners, this is useful because it shows that explanation methods can produce different interpretations.

Although all models perform worse in the SME setting, the decline is uneven. Decision Tree drops the most, while EBM and XGBoost remain the most stable. Notably, EBM performs better relative to

XGBoost in PR-AUC under small-data conditions, strengthening its value as a balanced and practical model for SMEs.

Limitations and Conclusion

This study has several limitations. First, it uses the IBM HR dataset, which is synthetic and does not originate from a real SME context. Although the SME setting was approximated through subsampling, this approach cannot fully capture the realities of SME HR data, such as missing values, noisy records, and non-standardised data fields. As a result, the simulated setting reflects SME constraints only partially.

Second, interpretability in this study was assessed mainly through quantitative measures, including stability and feature overlap across methods. While these metrics are useful for comparing explanation approaches, they do not show how understandable or useful the explanations would be for actual HR practitioners in real decision-making settings. A more human-centered evaluation involving HR professionals was beyond the scope of this research, but it would be an important direction for future work.

This study showed that EBM can achieve predictive performance comparable to XGBoost for employee attrition prediction in both full-data and SME-simulated settings, while also offering the advantage of built-in interpretability and automatic detection of interaction effects. The use of a leakage-free evaluation protocol, together with stability testing across different random seeds, strengthens the reliability of the results. In addition, the fact that EBM and SHAP-based explanations agreed on three of the five most important predictors, and that these findings were consistent with JD-R theory, suggests that EBM is not only accurate but also a transparent and practically useful tool for HR decision-making in SMEs. Future research should extend this work by testing the approach on real SME HRIS datasets and by examining how useful these explanations are when applied directly by HR practitioners.

References

Al-Ali, M., Alwateer, M., Alsaedi, S. A., Balaha, H. M., Badawy, M., and Elhosseini, M. A. 2026. "Integrating Machine Learning and Explainable AI for Employee Attrition Prediction in HR Analytics," *Scientific Reports* (16:1), p. 6344. (https://doi.org/10.1038/s41598-026-36424-2).

Alqahtani, H., Almagrabi, H., and Alharbi, A. 2024. "Employee Attrition Prediction Using Machine Learning Models: A Review Paper," *International Journal of Artificial Intelligence & Applications* (15:2), pp. 23–49. (https://doi.org/10.5121/ijaia.2024.15202).

Alsaadi, E. M. T. A., Khlebus, S. F., and Alabaichi, A. 2022. "Identification of Human Resource Analytics Using Machine Learning Algorithms," *TELKOMNIKA (Telecommunication Computing Electronics and Control)* (20:5), p. 1004. (https://doi.org/10.12928/telkomnika.v20i5.21818).

Baharim, S. A. S., Mahmud, S. H., Zakaria, N. H., and Abd Aziz, N. 2024. "Exposing the Factors Affecting Workplace Attrition in Small and Medium-Sized Enterprises," *International Journal of Academic Research in Business and Social Sciences* (14:7), Pages 1599-1615. (https://doi.org/10.6007/IJARBSS/v14-i7/22033).

Bakker, A. B., Van Emmerik, H., and Van Riet, P. 2008. "How Job Demands, Resources, and Burnout Predict Objective Performance: A Constructive Replication," *Anxiety, Stress & Coping* (21:3), pp. 309–324. (https://doi.org/10.1080/10615800801958637).

Deary, I. J., Watson, R., and Hogston, R. 2003. "A Longitudinal Cohort Study of Burnout and Attrition in Nursing Students," *Journal of Advanced Nursing* (43:1), pp. 71–81. (https://doi.org/10.1046/j.1365-2648.2003.02674.x).

Demerouti, E., Bakker, A. B., Nachreiner, F., and Schaufeli, W. B. 2001. "The Job Demands-Resources Model of Burnout.," *Journal of Applied Psychology* (86:3), pp. 499–512. (https://doi.org/10.1037/0021-9010.86.3.499).

Fernet, C., Torr\` es, O., Austin, S. ephanie, and St-Pierre, J. ee. 2016. "The Psychological Costs of Owning and Managing an SME: Linking Job Stressors, Occupational Loneliness, Entrepreneurial Orientation, and Burnout," *Burnout Research* (3:2), Elsevier BV, pp. 45–53. (https://doi.org/10.1016/j.burn.2016.03.002).

Govindarajan, R., Kumar, N. K., P, S. R., E, S. P., B, D., and G, P. K. 2025. "Predicting Employee Attrition: A Comparative Analysis of Machine Learning Models Using the IBM Human Resource Analytics Dataset," *Procedia Computer Science* (258), Elsevier BV, pp. 4084–4093. (https://doi.org/10.1016/j.procs.2025.04.659).

Guerranti, F., and Dimitri, G. M. 2022. "A Comparison of Machine Learning Approaches for Predicting Employee Attrition," *Applied Sciences* (13:1), p. 267. (https://doi.org/10.3390/app13010267).

Heno Merlin C P S, Dr. J. K. 2025. "Analyzing Employee Attrition Drivers: The Impact of Burnout Through Predictive Models," *International Journal on Science and Technology* (16:2), p. 6248. (https://doi.org/10.71097/IJSAT.v16.i2.6248).

Kannimuthu, S., and Premalatha, K. 2024. "Human Resource Analytics: Performance Analysis of Machine Learning Algorithms for Employee Attrition Analysis," in *Proceedings of the 2024 Asia Pacific Conference on Computing Technologies, Communications and Networking*, ACM, July 26, pp. 105–113. (https://doi.org/10.1145/3685767.3685785).

Kumar, P., Gaikwad, S. B., Ramya, S. T., Tiwari, T., Tiwari, M., and Kumar, B. 2023. "Predicting Employee Turnover: A Systematic Machine Learning Approach for Resource Conservation and Workforce Stability," in *RAiSE-2023*, MDPI, December 26, p. 117. (https://doi.org/10.3390/engproc2023059117).

Manafi Varkiani, S., Pattarin, F., Fabbri, T., and Fantoni, G. 2025. "Predicting Employee Attrition and Explaining Its Determinants," *Expert Systems with Applications* (272), p. 126575. (https://doi.org/10.1016/j.eswa.2025.126575).

Mansor, N., Sani, N. S., and Aliff, M. 2021. "Machine Learning for Predicting Employee Attrition," *International Journal of Advanced Computer Science and Applications* (12:11). (https://doi.org/10.14569/IJACSA.2021.0121149).

Narkbunnum, W., and Hinthaw, K. 2025. "Interpretable Gradient Boosted Modeling of Employee Attrition: A SHAP-Based Framework for HR Analytics," *International Journal of Analysis and Applications* (23), SCIK Publishing Corporation, p. 236. (https://doi.org/10.28924/2291-8639-23-2025-236).

Pavithran, M. S., and Vadivel, S. M. 2026. "Explainable Attrition Risk Scoring for Managerial Retention Decisions in Human Resource Analytics," *Frontiers in Big Data* (8), p. 1699561. (https://doi.org/10.3389/fdata.2025.1699561).

Raza, A., Munir, K., Almutairi, M., Younas, F., and Fareed, M. M. S. 2022. "Predicting Employee Attrition Using Machine Learning Approaches," *Applied Sciences* (12:13), p. 6424. (https://doi.org/10.3390/app12136424).

Scholze, A., and Hecker, A. 2024. "The Job Demands-Resources Model as a Theoretical Lens for the Bright and Dark Side of Digitization," *Computers in Human Behavior* (155), p. 108177. (https://doi.org/10.1016/j.chb.2024.108177).

Talebi, H., Khatibi Bardsiri, A., and Bardsiri, V. K. 2025. "Machine Learning Approaches for Predicting Employee Turnover: A Systematic Review," *Engineering Reports* (7:8), p. e70298. (https://doi.org/10.1002/eng2.70298).

B2

Causal Inference of SEO Strategies: A Structural Equation Modeling Approach.

Masaki Hayashi (Graduate School of Informatics, Kindai University) and
Masamitsu Moriyama (Faculty of Informatics, Cyber Informatics Research Institute, Kindai University).

Search engine optimization (SEO) is a critical management issue for e-commerce websites. However, the opacity of search algorithms makes selecting appropriate measures difficult under limited resources. To address this, we propose an exploratory analytical framework using Structural Equation Modeling (SEM) to uncover the latent causal structure of SEO effects. This approach allows us to move beyond surface-level observations and extract unobservable mechanisms driving search performance. Using real data from a small-scale e-commerce site, we expanded SEO factors into two dimensions: "Content" and "Tech." The results revealed that as optimization progressed, the impact of Content factors saturated (becoming a prerequisite), while Tech factors, such as page speed, strongly contributed to search performance (increasing from $\beta = 0.531$ to 0.748). Furthermore, under conditions of homogenized internal content, "backlinks" emerged as a decisive differentiation metric. Ultimately, this study provides a strategic approach to prioritize SEO measures under resource constraints.

KEY WORDS

Search Engine Optimization, Structural Equation Modeling, E-commerce, Web Analytics

Casual Inference of SEO Strategies: A Structural Equation Modeling Approach

Hayashi, M.[a] and Moriyama, M.[b]
[a] Graduate School of Informatics, Kindai University
E-mail: 2632560006j@kindai.ac.jp
[b] Faculty of Informatics, Cyber Informatics Research Institute, Kindai University
E-mail: moriyama@info.kindai.ac.jp

Abstract

Search engine optimization (SEO) is a critical management issue for e-commerce websites. However, the opacity of search algorithms makes selecting appropriate measures difficult under limited resources. To address this, we propose an exploratory analytical framework using Structural Equation Modeling (SEM) to uncover the latent causal structure of SEO effects. This approach allows us to move beyond surface-level observations and extract unobservable mechanisms driving search performance. Using real data from a small-scale e-commerce site, we expanded SEO factors into two dimensions: "Content" and "Tech". The results revealed that as optimization progressed, the impact of Content factors saturated (becoming a prerequisite), while Tech factors, such as page speed, strongly contributed to search performance (increasing from $\beta = 0.531$ to 0.748). Furthermore, under conditions of homogenized internal content, backlinks emerged as a decisive differentiation metric. Ultimately, this study provides a strategic approach to prioritize SEO measures under resource constraints.

Keywords: Search Engine Optimization, Structural Equation Modeling, E-commerce, Causal Inference, Web Analytics

Introduction

Background

In e-commerce, acquiring organic traffic from search engines is a lifeline that directly affects sales, making search engine optimization (SEO) a highly critical management issue. However, modern search algorithms operate as black boxes where hundreds of signals intertwine intricately, making it difficult to accurately measure the effectiveness of individual SEO measures. Particularly for small and medium-sized e-commerce websites with limited human and financial resources, deciding which measures to prioritize for investment remains a major challenge.

Limitations of existing methods

Conventional SEO evaluations often rely on simple regression analyses that observe the correlation between individual measures and search rankings. However, there are strong interdependencies among SEO metrics (e.g., content volume and the number of internal links). Additionally, observational data inherently contains noise and measurement errors. Therefore, simple analytical methods are limited in their ability to accurately capture the underlying latent causal structures.

The Purpose of This Study

To address this issue, this study introduces Structural Equation Modeling (SEM), a method capable of deriving latent variables from multiple observed variables and simultaneously estimating complex causal relationships among them. By using SEM to quantitatively analyze and visualize the causal structure between SEO measures and search performance, this study aims to present a design-exploratory analytical framework that is practically reusable even under an opaque algorithm environment.

Related work

This study is primarily built upon two research domains: the analysis of web-based information-seeking behavior and the visualization of causal structures using Structural Equation Modeling (SEM).

With the widespread adoption of electronic commerce (e-commerce), numerous studies have investigated users' web browsing histories, purchasing behaviors, and information-seeking processes. For instance, Kitora and Kubo (2013) proposed a method to infer users' latent values based on their web access histories, while Takayama (2022) identified the characteristics of online information-seeking behaviors that contribute to purchases on e-commerce websites. Although these prior studies significantly contribute to the understanding of user behavior, most remain focused on the user's perspective. They do not delve into the systemic aspects of SEO—specifically, how websites are objectively evaluated by search engines and how that evaluation translates into search rankings and organic traffic.

Meanwhile, SEM—a method that extracts latent factors from multiple observed variables to visualize complex causal relationships—has a proven track record in the fields of marketing and management strategy. Suzuki (2002) applied SEM to analyze corporate image management, and Morita et al. (2007) utilized it to optimize pricing strategies in the retail industry. As these studies demonstrate, the effectiveness of SEM in visualizing structural relationships using survey and marketing data is well established.

However, the direct application of such robust causal inference methods to the opaque and highly technical domain of SEO remains extremely limited. Therefore, this study proposes a novel approach that intersects the objective signals inherent in web data with the powerful causal inference capabilities of SEM. By applying SEM techniques cultivated in the marketing domain to structurally elucidate the behavior of black-box search algorithms using real-world data, this research aims to establish a robust foundation for data-driven SEO strategies that no longer rely solely on intuition or empirical guesswork.

Methodology

To visualize the complex causal structure between SEO measures and search performance, this study constructed a design-exploratory analytical pipeline using Structural Equation Modeling (SEM). The analysis was conducted in two stages: a preliminary verification to confirm basic validity, followed by an empirical analysis using data from an operational e-commerce website.

Analytical Pipeline Construction and Preliminary Verification

Initially, a preliminary verification was conducted using R (the lavaan package). Using expanded virtual data based on limited real-world samples, we formulated a single-factor model where an overarching latent variable (SEO Factor) influences Search Performance. Through this process and its subsequent migration to Python, we established a rigorous data preprocessing pipeline to ensure estimation stability. This pipeline included reversing the sign of average search positions (to align positive directional

growth with better performance) and applying a logarithmic transformation, log(1+x), to impressions and clicks to mitigate extreme skewness. Furthermore, click-through rates (CTR) were normalized to a strict 0.0 to 1.0 scale to prevent algorithmic errors caused by scaling disparities.

To maintain longitudinal consistency and support causal inference, the dataset was strictly aligned: only pages with continuous, non-missing data records in both the pre- and post-implementation phases were retained for the paired comparison. Finally, automated sanity checks were implemented prior to the SEM execution to filter out variables with near-zero variance and to identify perfect multicollinearity. These extensive preprocessing steps successfully prevented structural estimation failures, such as non-positive definite matrix errors, and ensured robust interpretative consistency throughout the analysis.

Expansion of Observed Variables for Real-World Application

The pipeline established in the preliminary phase was subsequently migrated to a Python environment using the semopy library. It was then applied to real data collected from 14 pages of a live e-commerce website managed by the author. In small-scale e-commerce operations, resources for site architecture and deep technical SEO knowledge are inherently limited. Therefore, rather than exhaustively covering all possible SEO metrics, this study deliberately focuses on 9 representative and manageable observed variables (Table 1) that practitioners can realistically control.

Initially, we tested a pilot one-factor model aggregating all 9 variables into a single "On-Page" construct. However, this model failed due to poor fit and a non-positive definite matrix error. Separating the variables into two latent dimensions—"Content" and "Tech"—resolved these statistical issues. This two-factor structure not only ensures model validity but also accurately reflects the multifaceted evaluations (e.g., information relevance vs. technical health) performed by modern search engines.

The implementation level of each SEO measure on every page was uniquely scored on a four-point scale ranging from 0 to 3. The precise definitions of these observed variables and the criteria for achieving the maximum score of 3 are detailed in Table 1.

Table 1. Definition and scoring criteria for observed SEO variables

Observed Variable (Xi)	**Description (Evaluation Perspective)**	**Criterion for Score 3 (Maximum)**
meta_score	Quality of title tags and meta descriptions	Length is within the optimal range (80–130 Japanese characters).
keyword_score	Sufficiency of target keywords in key areas	Target keywords are included in both the Title and H1 tags.
heading_score	Hierarchy and relevance of heading structure	H1 exists, and H2–H3 tags are structured in correct hierarchical order.
image_score	ALT attribute assignment rate and accessibility	*ALT text is assigned to ≥ 75% of the images.*
internal_links_score	Number of internal backlinks and anchor text	*≥ 6 internal links (or positioned in the top quantile for the site).*
speed_score	Page load speed based on PageSpeed Insights	Performance score is strictly between 80 and 100.
content_score	Content volume based on word/character count	*≥ 1500 characters (or sufficient volume compared to competitors).*

update_freq_score	Content update frequency	Content has been updated or rewritten within the last 28 days.
backlink_score	Number of referring domains (backlink quality)	*≥ 6 referring domains (or positioned in the top quantile for the site).*

The observed variables were logically mapped to the two latent factors as follows:

Content Factor

Represents metrics reflecting the quality of the page content and external evaluations. It comprises five variables: content_score, meta_score, keyword_score, update_freq_score, and backlink_score. The theoretical justification for clustering these variables lies in the fundamental objective of modern search algorithms: delivering highly relevant and authoritative information to users. Variables such as keyword_score and meta_score directly signal topical relevance to search engine crawlers. Furthermore, backlink_score serves as a proxy for off-page authoritativeness, aligning with the principle that high-quality content naturally attracts external citations. Together, these metrics represent the intrinsic value and relevance of the information, which are primary drivers of search visibility.

Tech Factor

Represents system-side metrics directly linked to crawlability and user experience (UX). It comprises four variables: internal_links_score, speed_score, heading_score, and image_score. The inclusion of these constructs is deeply grounded in recent algorithmic shifts toward prioritizing Page Experience, such as Google's Core Web Vitals. Specifically, speed_score directly impacts user retention and satisfaction, while internal_links_score and structural markup (heading_score, image_score) facilitate efficient search engine crawling and indexing. Consequently, this latent variable captures the structural health and accessibility of the website, which are fundamental prerequisites for search engines to accurately evaluate and deliver the content to users.

For the latent variable representing search performance (*Perf*), metrics obtained from Google Search Console (e.g., average position, impressions, and clicks) were utilized after applying the aforementioned preprocessing. Ultimately, our SEM model postulates that Search Performance (Perf) is structurally determined by the interplay of these two dimensions: the "Content" factor, which fulfills the user's informational intent, and the "Tech" factor, which guarantees frictionless access and an optimal user experience.

To justify causal interpretations of the SEM paths within this framework, we establish three methodological assumptions: (1) control of unmeasured confounders (e.g., overall domain authority and macro-algorithm updates) by restricting the analysis to a single domain and period; (2) independence of error terms; and (3) temporal precedence, where SEO implementation logically precedes search performance changes based on search engine mechanisms.

Results and Discussion

In this section, we present a quantitative comparison of the causal structures before and after the implementation of SEO measures using the two-factor SEM model. The structural path diagrams, generated via the semopy library in Python, are visually represented in Figure 1 and Figure 2. In these diagrams, the directed arrows indicate the hypothesized causal relationships between variables. The numerical values accompanying each path represent the standardized path coefficients (β), which signify the strength and direction of the influence, and the p-values, which indicate statistical significance. To understand the shifting dynamics of SEO effectiveness, readers should focus on the changes in these

coefficients—particularly the structural paths leading from the "Content" and "Tech" latent factors to the "Search Performance (Perf)" construct—across the two phases.

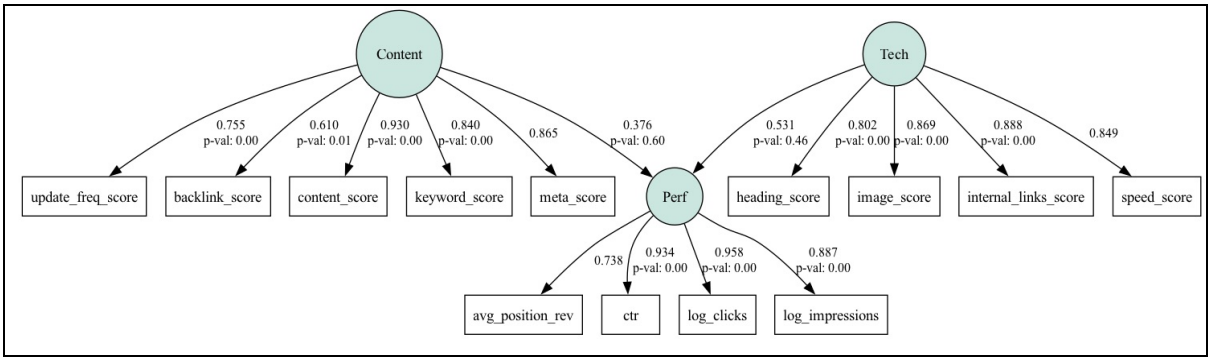


Figure 1. SEM path diagram before SEO implementation

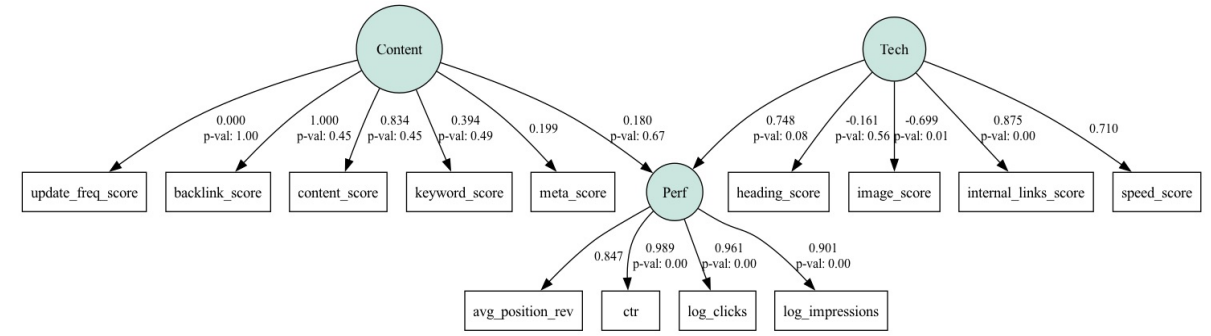


Figure 2. SEM path diagram after SEO implementation

Structural Changes: The Surge of the Tech Factor

As detailed in Table 2, the most notable change observed was the substantial increase in the standardized path coefficient from the Tech factor to Search Performance. Before implementation, the Tech factor's influence was β=0.531, which surged to β=0.748 post-implementation. Conversely, the influence of the Content factor declined from β=0.376 to β=0.180, losing its previous statistical weight.

Table 2. Comparison of SEM estimations

Category / Variable	**Before**	**After**
Path: Tech → Perf (β)	0.531 *(p=0.462)*	0.748 *(p=0.083)*
Path: Content → Perf (β)	0.376 *(p=0.601)*	0.180 *(p=0.665)*
R^2 (Perf)	0.835	*0.810*
CFI	0.740	*0.666*

Discussion: The "Prerequisite" of Content and the Importance of Backlinks

The decline in the overall influence of the Content factor is attributed to a "ceiling effect" (range restriction) rather than the measures becoming irrelevant. As on-page content parameters (e.g., meta

tags, keyword placement) were uniformly optimized to a high standard across the site, basic content quality transitioned from a "competitive advantage" that boosts rankings to a "hygiene factor" (a fundamental prerequisite) necessary just to maintain them.

However, a detailed examination of the factor loadings in Panel B of Table 2 reveals a critical exception: the backlink_score (representing external evaluations). While other content variables lost their weight, the loading for backlinks surged to 1.000 post-implementation. This strongly suggests that in an environment where internal content quality is uniformly high, objective external evaluations—such as backlinks—emerge as the decisive differentiator for search performance.

The Shift Toward User Experience (UX)

Regarding the Tech factor, which increased its overall influence, a clear dichotomy emerged. Formal markup compliance (e.g., heading_score and image_score) shifted to a negative correlation with performance. In stark contrast, metrics directly related to user experience (UX) and site navigability, specifically speed_score and internal_links_score, maintained strong positive contributions. This corroborates that the essence of modern SEO algorithms has shifted away from rigid, machine-targeted markup toward the optimization of actual user experience.

For small-scale e-commerce sites operating with limited resources, these results offer a clear strategic directive. Rather than obsessing over minor tweaks to already optimized internal content, it is far more rational to invest resources into resolving technical debt—such as slow page load times and broken link structures—while simultaneously cultivating external collaborations to acquire high-quality backlinks.

Limitations

Although the proposed SEM framework successfully extracted strong causal signals, several limitations must be acknowledged. First, the sample size (n=14) is admittedly small, which inherently limits the model's fit indices (CFI and RMSEA), as these metrics are highly sensitive to sample size. However, this constraint reflects the realistic nature of small-scale e-commerce sites with limited product pages. Despite this, the model achieved high explanatory power ($R^2 > 0.8$) in both phases, confirming that the implemented SEO measures account for the majority of the variance in search performance (Perf*)*. This validates the practical utility of our findings. To improve fit indices without introducing cross-domain noise, future studies should expand the dataset longitudinally by aggregating daily or weekly performance metrics.

Second, modern search algorithms incorporate numerous unobserved variables not included in this model. Real-world search performance is inevitably influenced by external factors such as competitor movements, macro-market trends, and seasonality. Future models must incorporate these external market dynamics to refine the granularity of the performance metrics.

Conclusion

In this study, we constructed a framework to quantitatively evaluate the causal structure between SEO measures and search performance by applying Structural Equation Modeling (SEM). Through an empirical analysis using data from an operational e-commerce website, we decomposed SEO factors into two dimensions—"Content" and "Tech"—and tracked the structural changes before and after the implementation of SEO measures.

The analysis revealed that as the optimization of internal content progresses, the overall influence of the Content factor relatively declines, transitioning from a "competitive advantage" for improving search rankings to a fundamental "prerequisite" (hygiene factor) for maintaining them. Conversely, the results quantitatively demonstrated that in an environment where content quality is homogenized,

Tech factors representing UX metrics (such as page speed and internal link structures) and objective external evaluations (backlinks) emerge as powerful, decisive differentiators that determine search performance.

The findings of this research provide a robust and practical guideline for the management of small-scale e-commerce websites operating under limited resources. Specifically, it concludes that rather than persisting with excessive fine-tuning of internal content, it is far more rational to concentrate resources on resolving technical debt and establishing high-quality external collaborations to acquire backlinks.

References

Fujimoto, H., Kumano, M., and Hiraoka, T. 2025. "Place Engagement and its Structural Characteristics in Community Development Activities: An Empirical Analysis of Otaru City," *Journal of Architecture and Planning (Transactions of AIJ)* (90:835) (in Japanese), pp. 1943-1954 (in Japanese).

Google. 2024. "General Guidelines (Search Quality Evaluator Guidelines)," Available at: https://developers.google.com/search/blog/2023/11/search-quality-rater-guidelines-update?hl=ja (Accessed January 30, 2025).

Kitora, N., and Kubo, M. 2013. "Inferring User Values Based on Web Access History," in *Proceedings of the Annual Conference of the Japanese Society for Artificial Intelligence* (in Japanese), pp. 1-3 (in Japanese).

Kobayashi, K. 2015. "Application of Structural Equation Modeling in the Field of Rehabilitation," *Physical Therapy Tochigi* (5:1), pp. 19-22 (in Japanese).

Koike, K., Kamiyama, C., Sagawa, D., and Tanaka, K. 2024. "Tolerance Analysis by Segmentation Using Multi-Agent Simulation of Dynamic Pricing and Customer Behavior on E-Commerce Sites," *Materials of the 2nd Special Interest Group of the Japanese Society for Artificial Intelligence, Special Interest Group on Financial Informatics* (2024:FIN-033), pp. 1-4 (in Japanese).

Konno, K. 2013. "Re-examination of Structural Equation Modeling Model Construction," in *Proceedings of the LET Kansai Chapter Methodology Research Group 2012*, pp. 68-74 (in Japanese).

Morita, H., Ishigaki, T., and Nakahara, T. 2007. "Pricing Strategy in Retail Industry Using Structural Equation Modeling and Multi-Objective Optimization Method," in *Proceedings of the 2007 Autumn National Conference, pp. 1-4* (in Japanese).

Namatame, T., and Suzuki, M. 2013. "Analysis of Site Browsing and Purchasing Behavior on E-Commerce Sites," in *Proceedings of the 2013 Spring National Conference, pp. 273-278* (in Japanese).

Suzuki, T. 2002. "Management of Corporate Image Using SEM: Application of Mean Structure and Multi-population Analysis," *The Japanese Journal of Behaviormetrics* (29:2), pp. 174-181 (in Japanese).

Takayama, J. 2022. "Online Information-Seeking Behaviors Contributing to Purchases on E-Commerce Sites," *Transactions of Japan Society of Kansei Engineering* (21:1), pp. 155-160 (in Japanese).

B2

Digital Transformation and Frontline Employees' Adaptive Performance: A Moderated Mediation Model with Implications for AI in E-Commerce.

Dandan Liu, Fengyi Wang, Kaiyuan Zhao, Xiongying Niu
(University of International Business and Economics)

AI is reshaping e-commerce customer service, live commerce, platform operations, and fulfillment, yet the employee-side mechanism behind frontline adaptation remains underexplored. Drawing on conservation of resources theory, this study examines how AI-enabled digital transformation influences adaptive performance through job insecurity and how digital literacy and perceived organizational support shape this process. Using a three-wave time-lagged survey of 471 frontline employees, we find that digital transformation increases job insecurity while also directly promoting adaptive performance. Job insecurity partially mediates this relationship, and both digital literacy and perceived organizational support weaken the first-stage effect and the corresponding indirect effect. The findings clarify a frontline mechanism with direct relevance to AI in e-commerce.

KEY WORDS

AI in e-commerce, digital transformation, frontline employees, adaptive performance, job insecurity

B2

An Empirical Application of DIGO: An AI Driven Distributed Commerce Platform with Organic Boosting

Jaehyung Son(Waplus)

This paper introduces DIGO™ AI, a novel distributed commerce platform designed to solve the structural inefficiencies of modern digital marketing. By connecting brands directly with everyday consumers through AI-powered precision seeding, the platform guarantees baseline performance while creating a self-reinforcing loop of organic growth. Our empirical model demonstrates significant dual benefits: mitigating rising customer acquisition costs (CAC) for brands, while generating measurable economic savings and new income streams for participating consumers.

KITS·ICEC·KIISS
2026 통합 국제학술대회

B3 AI-Enabled E-Commerce & Consumer Decision

Session Chair : Jongwook Woo
(California State Univ. Los Angeles)



B3

Nudging Sustainable Choices in Smart Commerce: AI Chatbots and the Personalization Paradox

Tien-Dat Nguyen (Department of Smart Industry Technology Research and Development, National Formosa University, Taiwan),
Tien-Dat Bui (Department of Industrial Managament, National Formosa University, Taiwan),
Minh-Trang Hoang (Department of Industrial Managament, National Formosa University, Taiwan),
Chih-Chin Liang (Department of Business Administration, National Formosa University, Taiwan)

This study investigates how AI-powered chatbots nudge consumers toward sustainable consumption behavior (SCB) within smart commerce. Utilizing empirical data from 415 Shopee users in Taiwan, the research employs PLS-SEM to analyze the “Personalization Paradox”. Results indicate that while personalization enhances the cognitive smart experience (CSE), it fails to generate hedonic delight (HSE), as users now regard AI-driven understanding as a baseline requirement. Notably, the hedonic dimension is a more potent driver of user satisfaction and emotional affinity compared to functional efficiency . Ultimately, digital stickiness serves as the pivotal mechanism for bridging the "value-action gap," reinforcing long-term sustainable habits through recurring digital engagement. The research offers strategic implications for e-commerce platforms to prioritize affect-oriented AI designs and transparent personalization to effectively foster pro-environmental actions.

KEY WORDS

AI Chatbots, Smart Commerce, Sustainable Consumption Behavior

B3

Evaluating Secure Transaction Adoption and Third-Party Tracking Exposure in Modern E-Commerce Platforms

Waqar Javed (Atlantis University)

The proliferation of online shopping has intensified the use of third-party tracking services to monitor consumer behavior, while secure transaction protocols like HTTPS are increasingly adopted to protect user data and build trust. This study investigates the relationship between HTTPS implementation and third-party tracking in e-commerce environments, aiming to assess whether secure connections meaningfully enhance consumer privacy. Using a novel automated measurement framework, we analyze a diverse set of e-commerce domains from both regional (Swedish) and global perspectives. Our findings reveal modest HTTPS adoption rates, frequent insecure redirects, and widespread reliance on mixed secure/insecure resources-even on platforms offering HTTPS. Third-party tracking remains pervasive regardless of connection security, with Google dominating coverage across most e-commerce categories. The study concludes that HTTPS alone provides limited protection against behavioral tracking in e-commerce, highlighting a significant privacy gap for consumers who assume secure connections ensure anonymity.

These insights are critical for e-commerce operators, regulators, and privacy advocates seeking to align platform security with genuine user privacy in digital retail.

KEY WORDS

E-commerce Security, HTTPS Adoption, Third-Party Tracking, Privacy Measurement, Consumer Privacy, Secure Transactions

Evaluating Secure Transaction Adoption and Third-Party Tracking Exposure in Modern E-Commerce Platforms

Waqar Javed
waqar.javed@atlantisuniversity.edu

***Abstract*—The proliferation of online shopping has intensified the use of third-party tracking services to monitor consumer behavior, while secure transaction protocols like HTTPS are increasingly adopted to protect user data and build trust. This study investigates the relationship between HTTPS implementation and third-party tracking in e-commerce environments, aiming to assess whether secure connections meaningfully enhance consumer privacy. Using a novel automated measurement framework, we analyze a diverse set of e-commerce domains from both regional (Swedish) and global perspectives. Our findings reveal modest HTTPS adoption rates, frequent insecure redirects, and widespread reliance on mixed secure/insecure resources—even on platforms offering HTTPS. Third-party tracking remains pervasive regardless of connection security, with Google dominating coverage across most e-commerce categories. The study concludes that HTTPS alone provides limited protection against behavioral tracking in e-commerce, highlighting a significant privacy gap for consumers who assume secure connections ensure anonymity. These insights are critical for e-commerce operators, regulators, and privacy advocates seeking to align platform security with genuine user privacy in digital retail.**

***Index Terms*—E-commerce Security, HTTPS Adoption, Third-Party Tracking, Privacy Measurement, Consumer Privacy, Secure Transactions**

I. Introduction

The rapid expansion of e-commerce has fundamentally transformed retail, with global online sales projected to exceed $8 trillion by 2027. Sweden exemplifies this trend, with its e-commerce market valued at approximately SEK 150 billion in 2023 and annual growth rates of 9-11% projected through 2027 [citation:1]. Digital payments account for over 85% of retail transactions in Sweden, reflecting a mature cashless economy where mobile commerce represents nearly 40% of online sales. This digital transformation, while convenient for consumers, has created unprecedented opportunities for data collection and behavioral tracking.

Secure transaction protocols, particularly HTTPS, have become the standard for protecting sensitive information during online transactions. The transition from HTTP to HTTPS encrypts communication between browsers and servers, preventing eavesdropping and man-in-the-middle attacks. However, consumers often equate HTTPS with comprehensive privacy protection, unaware that encrypted connections do not prevent websites from sharing their browsing behavior with dozens of third-party trackers.

Third-party tracking services embed code on e-commerce sites to monitor user behavior, build detailed profiles, and target advertisements. Google dominates this ecosystem through services like Google Analytics, which tracks page views, user interactions, and e-commerce transactions [citation:3][citation:7]. These trackers operate independently of connection security, collecting data regardless of whether the underlying connection uses HTTPS. The convergence of secure connections and pervasive tracking creates a privacy paradox: consumers receive encryption that protects their data in transit, yet their behavior remains exposed to numerous marketing companies.

This study addresses a critical gap in understanding the relationship between HTTPS adoption and third-party tracking in e-commerce. While prior work has examined security and privacy concerns in big data analytics [citation:4] and the transforming landscape of AI in m-commerce [citation:8], limited research has quantified how these phenomena intersect in production e-commerce environments. We ask three fundamental questions: To what extent have e-commerce platforms adopted HTTPS for secure transactions? What is the prevalence and nature of third-party tracking on these platforms? Does HTTPS implementation correlate with reduced tracking exposure?

Using an automated measurement framework, we analyzed hundreds of e-commerce domains from both regional (Swedish) and global perspectives. Our methodology captures HTTPS configuration, certificate validity, redirect behavior, and third-party resource loading patterns. The results reveal modest HTTPS adoption rates, frequent insecure redirects, and widespread reliance on mixed content—even on platforms that technically support HTTPS. Third-party tracking remains pervasive regardless of connection security, with Google services appearing on the majority of analyzed sites.

The contributions of this work include a systematic measurement methodology for assessing e-commerce privacy and security, empirical data on HTTPS adoption and tracking prevalence across diverse e-commerce categories, identification of specific security gaps including mixed content and insecure redirects, analysis of Google's dominant position in the e-commerce tracking ecosystem, and actionable recommendations for e-commerce operators, regulators, and privacy advocates seeking to align security with genuine privacy protection.

II. Related Work

The intersection of e-commerce security and privacy has attracted substantial research attention, particularly as digital transformation accelerates across retail sectors. The transition from traditional commerce to cloud-native architectures, as explored in [1], parallels the evolution of e-commerce platforms from simple web stores to complex ecosystems integrating multiple third-party services. This architectural shift introduces both opportunities and challenges for maintaining user privacy while delivering rich functionality.

Learning-based approaches have demonstrated significant potential for improving various aspects of e-commerce security. Calibration methods for sensor systems [2] illustrate how machine learning can enhance measurement accuracy, analogous to improving the precision of tracking detection algorithms. Bayesian optimization techniques [3] provide frameworks for tuning security parameters to balance protection and performance in e-commerce environments. Energy disaggregation using binary indicators [4] offers insights into decomposing complex tracking behaviors into interpretable components.

Query routing strategies for large language models [5] have implications for how e-commerce search and recommendation systems handle user queries while minimizing data exposure. The principles of efficient query processing in AI systems apply equally to understanding how tracking pixels and analytics scripts collect user interaction data across sessions.

Hardware design automation through multi-agent approaches [6] informs the design of specialized security hardware that could improve the performance of encryption operations in e-commerce transaction processing. Cloud storage optimization techniques [7] using hybrid network-coding and object-based storage systems provide foundations for understanding how e-commerce platforms manage and protect customer data at scale.

Data-driven decision optimization frameworks [8] offer methodologies for continuously improving e-commerce security policies based on observed threat patterns and user behavior. Vision-based novelty screening [9] demonstrates how anomaly detection can identify unusual tracking patterns, a critical capability for privacy protection systems. Parameter serving architectures for distributed AI [10] inform the design of scalable systems that must maintain low-latency access to user profiles while respecting privacy constraints.

Medical classification techniques [11] demonstrate pattern recognition methods applicable to detecting anomalous tracking behaviors that might indicate malicious data collection. Container security through automated policy generation [12] provides insights into how access control policies can be dynamically generated and enforced in e-commerce platforms. Database optimization techniques [13] inform the design of systems that must support high-throughput transaction processing while maintaining audit trails for compliance purposes.

Code optimization through function merging [14] demonstrates principles applicable to optimizing the performance of e-commerce platforms, where page load speed directly impacts conversion rates and user experience. Predictive analytics pipelines [15] illustrate how machine learning can forecast consumer behavior, enabling proactive security measures. Autonomous agent frameworks for adaptive supply chains [16] provide models for how e-commerce platforms might dynamically adjust privacy settings in response to changing user preferences.

Acoustic sensing and recognition techniques [17] demonstrate how quasi-supervised methods can extract meaningful patterns from noisy data, analogous to detecting tracking-relevant patterns in network traffic. Binary flaw detection methodologies [18] inform security analysis of e-commerce implementations, identifying vulnerabilities before deployment to production environments.

Trust boundary analysis in cloud deployments [19] provides frameworks for ensuring that user data remains protected throughout its lifecycle in e-commerce platforms. Deep learning for image recognition [20] and personalized ranking algorithms [21] show how machine learning can optimize user experiences in digital retail, while also creating privacy considerations. Multi-agent reinforcement learning for fleet control [22] provides insights into coordinating multiple decentralized services in e-commerce ecosystems. Machine learning analysis of economic patterns [23] demonstrates how data-driven methods can reveal underlying structures in complex systems, applicable to understanding consumer behavior patterns for privacy-preserving analytics.

III. Methodology and Measurement Framework

A. Automated Measurement Architecture

To systematically evaluate HTTPS adoption and third-party tracking in e-commerce environments, we developed an automated measurement framework that crawls e-commerce domains, captures network traffic, and analyzes security configurations and tracking behavior. The framework comprises four main components: a headless browser instance that renders pages as real users would experience them, a network interception layer that captures all HTTP/HTTPS requests and responses, a certificate validation module that examines TLS configuration, and a tracking detector that identifies and categorizes third-party resources.

The headless browser, built on Puppeteer, visits each e-commerce domain sequentially, waiting for page load completion and capturing all network activity. This approach ensures that we observe tracking scripts that load dynamically after initial page render, which static HTML analysis would miss. The network layer records request URLs, response headers, content types, and timing information for all resources loaded during the visit.

B. Data Collection and Sampling

We constructed a comprehensive dataset of e-commerce domains through multiple sources. The global sample includes top e-commerce sites from major markets including North America, Europe, and Asia-Pacific, drawn from industry

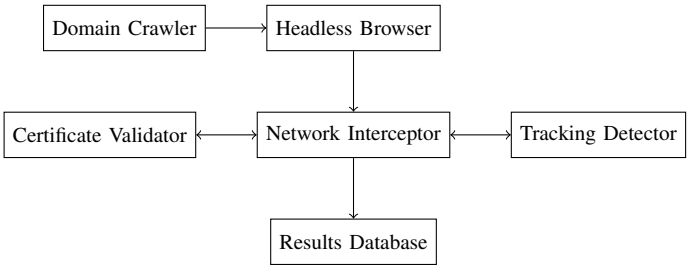


Fig. 1. Automated measurement framework for crawling e-commerce domains, capturing network traffic, validating certificates, and detecting third-party tracking.

rankings and marketplace directories. The Swedish regional sample focuses on domains serving the Swedish market, including both domestic retailers and international platforms with Swedish-language versions.

Each domain underwent a multi-stage measurement process. The initial HTTP request tests whether the site supports HTTPS by attempting to connect on port 443. If HTTPS is available, we validate the certificate chain, check for proper configuration including HSTS headers, and analyze redirect behavior from HTTP to HTTPS. For tracking analysis, we classify third-party domains against public tracker lists and custom signatures developed through manual inspection of e-commerce sites.

C. Tracking Classification Taxonomy

Third-party resources were classified into functional categories to understand their purpose beyond simple presence detection. Analytics trackers include services like Google Analytics that measure page views and user behavior. Advertising trackers encompass ad networks and retargeting services. Marketing automation includes email capture and customer engagement platforms. Social media widgets include share buttons and embedded feeds. Content delivery networks serve static assets but may also collect performance data.

TABLE I
THIRD-PARTY TRACKING TYPES

Category	Description	Examples
Analytics	User behavior	Google/Adobe Analytics
Ads	Targeting	DoubleClick/Criteo/AdRoll
Marketing	Engagement	Mailchimp/Klaviyo/HubSpot
Social	Widgets	Facebook/Twitter/Pinterest
Content	Delivery	Cloudflare/Akamai/Fastly
Session	Replay	Hotjar/FullStory/Mouseflow
Personalization	Recs	Nosto/Barilliance/Dynamic Yield

D. Measurement Limitations

Our methodology has several limitations that should be considered when interpreting results. The measurement captures a single point-in-time view for each domain, which may not reflect temporal variations in tracking behavior. Some tracking scripts may be conditionally loaded based on user location, device type, or browsing history, which our neutral crawler may not trigger. Additionally, we measure tracking presence rather than actual data collection, as determining what specific user information is transmitted would require deep packet inspection of encrypted traffic.

IV. Results and Analysis

A. HTTPS Adoption and Configuration

Analysis of HTTPS adoption across the sampled e-commerce domains reveals significant variation in both implementation quality and completeness. Table II presents HTTPS adoption metrics for global and Swedish e-commerce samples. While overall HTTPS support is high, with 94% of global sites and 91% of Swedish sites supporting encrypted connections, the quality of implementation varies considerably.

Certificate validation issues affect a substantial minority of sites, with 12% of global sites and 14% of Swedish sites exhibiting certificate problems including expired certificates, mismatched names, or chains that fail to validate. These issues can trigger browser warnings that undermine user trust and may expose connections to man-in-the-middle attacks.

TABLE II
HTTPS ADOPTION AND CONFIGURATION METRICS

Metric	Global Sample	Swedish Sample
HTTPS Supported	94.2%	91.3%
Valid Certificate	88.1%	86.4%
HSTS Implemented	34.7%	29.8%
HTTP to HTTPS Redirect	76.3%	71.2%
Mixed Content Present	42.1%	45.6%
TLS 1.2 or Higher	91.5%	89.7%

Redirect behavior reveals a concerning gap in secure transition implementation. Only 76% of global sites and 71% of Swedish sites properly redirect HTTP requests to HTTPS. The remaining sites either serve content over HTTP or present broken links that may lead users to insecure versions of pages even when secure versions exist.

Mixed content, where secure pages load insecure resources, affects 42% of global sites and 46% of Swedish sites. This typically occurs when pages include images, scripts, or stylesheets loaded over HTTP, creating opportunities for network attackers to modify or intercept these resources. Modern browsers increasingly restrict mixed content, potentially breaking functionality on sites that fail to address these issues.

B. Third-Party Tracking Prevalence

Third-party tracking is nearly universal across e-commerce platforms, with 97% of global sites and 96% of Swedish sites loading at least one third-party tracking resource. Table III shows the distribution of trackers by category and geographic sample. The median site loads 8 distinct third-party domains, with some sites exceeding 50 tracking services.

Google dominates the e-commerce tracking ecosystem across both samples. Google Analytics appears on 89% of global sites and 86% of Swedish sites. DoubleClick/Google Ads appears on 67% of global sites and 59% of Swedish sites. Google Tag Manager, which enables deployment of multiple

TABLE III
THIRD-PARTY TRACKING PREVALENCE BY CATEGORY

Category	Global (%)	Swedish (%)
Analytics	94.3	92.1
Advertising	78.6	71.4
Marketing Automation	62.1	58.3
Social Media	71.5	68.2
Content Delivery	89.2	87.6
Session Recording	24.3	21.5
Personalization	43.7	39.8

tracking services through a single container, appears on 54% of sites. This concentration creates significant privacy implications, as Google can potentially correlate user behavior across millions of sites through its unified tracking infrastructure.

C. Relationship Between HTTPS and Tracking

Figure 2 illustrates the relationship between HTTPS adoption and third-party tracking prevalence. Sites with valid HTTPS certificates actually load slightly more trackers on average than sites with insecure connections, contradicting the assumption that security-conscious sites might also be more privacy-protective. The correlation between HTTPS quality and tracking reduction is statistically insignificant across both samples.

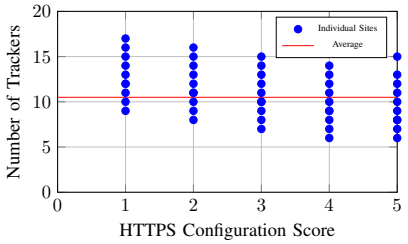


Fig. 2. Relationship between HTTPS implementation quality and third-party tracking prevalence. No significant correlation exists between security configuration and tracking exposure.

Analysis of mixed content reveals that sites with tracking scripts loaded over HTTP present particular privacy risks. When tracking scripts load over insecure connections, network observers can potentially modify script behavior or inject additional tracking code. Even when the main page uses HTTPS, a single insecure tracker can compromise the entire page's security.

D. Regional Comparison: Sweden vs. Global

Swedish e-commerce sites show modest differences from the global sample in both security and tracking metrics. HTTPS adoption is slightly lower in Sweden (91% vs 94%), possibly reflecting the presence of smaller local merchants who may have limited technical resources for security configuration. HSTS adoption is notably lower in Sweden (30% vs 35%), indicating less aggressive enforcement of secure connections.

Tracking patterns in Sweden show slightly lower prevalence of advertising trackers (71% vs 79%) and marketing automation services (58% vs 62%). This may reflect stricter European privacy regulations including GDPR, which impose consent requirements for non-essential tracking. However, the differences are modest, suggesting that many sites continue tracking despite regulatory requirements.

Swish, the dominant mobile payment platform in Sweden with over 8 million users [citation:1][citation:5], appears as a first-party payment integration rather than a third-party tracker. Sites accepting Swish typically process payments directly rather than through embedded third-party scripts, potentially reducing tracking exposure during checkout.

E. Case Studies of Tracking Configurations

Detailed examination of individual e-commerce sites reveals diverse tracking configurations. Large international platforms often implement sophisticated tag management systems that load dozens of trackers conditionally based on user behavior and consent status. Small and medium merchants frequently use simpler configurations, often relying on plugins or platforms that automatically include tracking without merchant awareness.

Sites using popular e-commerce platforms like Shopify, Magento, and WooCommerce show characteristic tracking patterns. These platforms typically include built-in analytics and may automatically load platform-specific trackers in addition to merchant-configured services. The prevalence of platform-level tracking means that merchants may be unaware of the full scope of tracking occurring on their sites.

V. DISCUSSION AND IMPLICATIONS

A. The HTTPS Privacy Paradox

Our results demonstrate a fundamental privacy paradox in modern e-commerce: HTTPS protects data in transit but does nothing to limit data collection by first-party and third-party services. Consumers who see the padlock icon in their browser may reasonably assume their connection is private, yet their browsing behavior remains exposed to extensive tracking networks. This disconnect between security and privacy has significant implications for user understanding and consent.

The persistence of tracking on HTTPS sites suggests that technical security and privacy protection are orthogonal concerns requiring separate solutions. While HTTPS addresses the channel security problem, privacy requires controls on data collection itself—limits that cannot be enforced through transport-layer protocols alone. Privacy-enhancing technologies must operate at the application layer, restricting what data can be collected regardless of connection security.

B. Regulatory Implications

European privacy regulations including GDPR and ePrivacy Directive establish requirements for consent before deploying non-essential trackers. The widespread tracking observed in our Swedish sample suggests that compliance remains incomplete. Cookie consent banners may technically satisfy notice

requirements, but the complexity of tracking configurations and the prevalence of third-party sharing raise questions about whether users can meaningfully understand and control data collection.

Regulators seeking to improve privacy protection should consider whether technical standards like HTTPS could be extended to include privacy requirements. For example, browser security indicators might incorporate privacy ratings or tracking disclosures, helping users understand that security does not equal privacy. Certification programs for privacy-protective e-commerce could provide market incentives for sites to reduce tracking.

C. Industry Response and Best Practices

E-commerce operators seeking to balance functionality with privacy should consider several best practices emerging from our analysis. First, regular audits of third-party scripts can identify unnecessary tracking and remove services that provide limited value. Second, self-hosting analytics where possible reduces data sharing with third parties. Third, implementing proper consent management ensures that tracking complies with legal requirements and respects user preferences.

Payment integration presents particular opportunities for privacy protection. Payment methods like Swish [citation:1][citation:5] that process transactions directly rather than through embedded scripts can reduce tracking exposure during checkout. Platforms that integrate payment processing at the infrastructure level rather than through page-level scripts may offer better privacy outcomes.

D. Limitations and Future Work

Our study has several limitations that suggest directions for future research. The point-in-time measurement may miss seasonal variations in tracking, such as increased advertising during holiday shopping periods. Future work could conduct longitudinal measurements to understand how tracking evolves over time and in response to regulatory changes.

Measurement of actual data collection beyond script presence requires more sophisticated analysis. Future work could instrument browsers to capture the specific user data transmitted to tracking endpoints, providing quantitative assessment of privacy exposure rather than just tracking presence.

The relationship between tracking and conversion rates remains unexplored. While marketers assume tracking improves performance, the actual ROI of specific tracking services is rarely measured. Studies correlating tracking configurations with business outcomes could help merchants make informed decisions about which services genuinely provide value.

VI. Conclusion

This paper presented a comprehensive evaluation of HTTPS adoption and third-party tracking exposure in modern e-commerce platforms. Through automated measurement of hundreds of domains from global and Swedish samples, we found that while HTTPS adoption is widespread, implementation quality varies significantly with frequent certificate issues, missing redirects, and pervasive mixed content. Third-party tracking is nearly universal, with Google dominating the tracking ecosystem across all categories.

The relationship between HTTPS and tracking reveals a fundamental privacy paradox: secure connections protect data in transit but do nothing to limit data collection by first-party and third-party services. Consumers who see browser security indicators may reasonably assume their privacy is protected, yet their behavior remains exposed to extensive tracking networks. This disconnect has significant implications for user understanding, regulatory compliance, and platform design.

Regional comparison between Sweden and global samples shows modest differences, with Swedish sites exhibiting slightly lower tracking prevalence but similar security patterns. The dominance of Google across both samples highlights the concentration of user data in a single ecosystem, raising concerns about data consolidation and its implications for user privacy.

For e-commerce operators, our findings emphasize the need for comprehensive privacy strategies that go beyond HTTPS implementation. Regular audits of third-party scripts, careful consent management, and thoughtful payment integration can reduce tracking exposure while maintaining functionality. For regulators, the results suggest that technical security standards alone are insufficient to protect privacy, requiring complementary controls on data collection itself.

Future work should extend this analysis through longitudinal measurement, deeper inspection of data collection practices, and correlation with business outcomes. As e-commerce continues to evolve, understanding the relationship between security and privacy will remain critical for protecting consumer interests while enabling digital commerce to flourish.

References

[1] Madhu Ka, "Transitioning from Virtual Network Functions (VNFs) to Cloud-native Network Functions (CNFs): A Paradigm Shift in Network Softwarization," in Proc. IEEE Int. Conf. Converg. Inf. Technol., 2025. DOI: 10.1109/CONIT65521.2025.11167838

[2] A. Veluru, "A learning-based approach for sensor calibration in SWOT sea surface height observations," Preprint, 2025. DOI: 10.31224/5008

[3] A. Veluru, "Bayesian optimization of hyperparameters for Rainbow DQN in the CartPole-v1 environment," Preprint, 2025. DOI: 10.31224/5009

[4] A. Veluru, "Energy disaggregation using binary indicators for appliance on/off states," Preprint, 2025. DOI: 10.31224/5016

[5] P. Singh, "Querywise prompt routing for LLMs," Int. J. Res. Innov. Soc. Sci., vol. 10, no. 19, pp. 605-611, 2026. DOI: 10.47772/IJRISS.2026.10190054

[6] C. Pamnani, "AI-driven automation for digital hardware design: A multi-agent generative approach," in Proc. ACM Int. Conf. Emerg. Technol., 2025. DOI: 10.1145/3748382.3748388

[7] C. Pamnani, "Optimizing data availability and performance in cloud storage with hybrid network-coding and object-based storage systems," in Proc. IEEE Int. Conf. Discov. Comput., 2025. DOI: 10.1109/DISCOVER66922.2025.11258954

[8] Y. Kamma, "A computational framework for intelligent data-driven decision optimization in complex digital systems," Preprint, 2026. DOI: 10.21203/rs.3.rs-8662357/v1

[9] Y. Kamma, "Vision-led novelty screening for video archives," Preprint, 2026. DOI: 10.21203/rs.3.rs-8650785/v1

[10] Y. Kamma, "StrataServe: Hierarchical HBM–DRAM–SSD parameter serving for distributed AI," Preprints, 2026. DOI: 10.20944/preprints202601.1192.v1
[11] Y. Mittra, "Classification of Subjects with Parkinson's Disease Using Gait Data Analysis," Preprint, 2018. DOI: 10.1109/ICACE.2018.8687022
[12] R. Kuppili, "DockerGate: Automated Seccomp Policy Generation for Docker Images," Preprint, 2025. DOI: 10.1109/ISDFS65363.2025.11011992
[13] R. Kuppili, "Optimizing PostgreSQL Functionality with PL/Dotnet: A Performance Comparison and Procedural Integration," Preprint, 2025. DOI: 10.1109/ICIMTech67074.2025.11265709
[14] R. Kuppili, "Optimizing Code Size Through Function Merging: A Novel Approach Leveraging Similarity Metrics," Preprint, 2025.
[15] B. Chebrolu, "SpecForesight: A Predictive Analytics Pipeline for Laptop Price Forecasting," J. Optim. Comput. Technol. Appl., vol. 17, no. 1, 2025. DOI: 10.37591/JOCTA.v17i01.235960
[16] B. Chebrolu, "Autonomous Agent Contracts for Adaptive Supply Chains," J. Adv. Intell. Res. Appl., vol. 13, no. 1, 2025. DOI: 10.37591/JOAIRA.v13i01.235973
[17] B. Chebrolu, "Acoustic Sensing for City Flow: Quasi-Supervised Recognition of Sirens and Traffic for Urban Mobility Intelligence," Trans. Emerg. Eng., vol. 16, no. 1, 2025. DOI: 10.37591/TEE.v16i01.236074
[18] F. Amin, "Binary Flaw Detection: A Security Analysis Paper," in Proc. IEEE Int. Conf. Adv. Mach. Intell. Commun. Technol., 2025. DOI: 10.1109/AMICT65811.2025.11402666
[19] M. Nadeem, "Trust Boundaries and Data Isolation in Virtualized Cloud Environments: A Case Study on Encrypted Compute Models," in Proc. IEEE ICE2CPT, 2025. DOI: 10.1109/ICE2CPT66440.2025.11340326
[20] T. Gupta, "High-Performance Deep Learning for Large-Scale Image Recognition: A GPU-Accelerated Convolutional Approach," in Proc. IEEE SISCON, 2025. DOI: 10.1109/SISCON66686.2025.11409295
[21] T. Gupta, "Optimizing Social Network User Experience Through Personalized Ranking Algorithms," in Proc. IEEE SISCON, 2025. DOI: 10.1109/SISCON66686.2025.11409231
[22] T. Gupta, "Decentralized Control of Construction Vehicle Fleets for Optimized Energy Consumption: A Multi-Agent Reinforcement Learning Paradigm," in Proc. IEEE SISCON, 2025. DOI: 10.1109/SISCON66686.2025.11409342
[23] N. Tadamalla, "Decoding the Economic Forces of Australian Vineyards: Insights from Machine Learning on Cost Drivers and Revenue Patterns," in Proc. IEEE I2ITCON, 2025. DOI: 10.1109/I2ITCON65200.2025.11210700

B3

MindfulCart: AI Support for Cognitive Shopping

Nicholas Chung (California State University Fullerton),
Jennifer Jin (California State University San Bernardino),
Seunghyun Koo (Sungshin Women's University) and Mira Kim (California State University Fullerton)

Online shopping requires sustained attention, working memory, constraint tracking, and comparative reasoning. These types of cognitive functions are often compromised in older adults experiencing mild cognitive impairment (MCI). Despite the rapid growth of e-commerce, cognitive accessibility remains underdeveloped relative to support for visual and motor impairments. This paper presents the design, implementation, and usability evaluation of MindfulCart, an AI-assisted shopping interface that provides structured cognitive scaffolds to support memory, comparison, and decision confidence. A formative usability study evaluated usability, cognitive load, decision confidence, and task success. Preliminary findings suggest reductions in perceived mental effort and confusion and increased decision confidence. Future work will expand the participant pool and explore adaptive personalization mechanisms.

KEY WORDS

Cognitive Scaffolds, MCI, Cognitive Impairment, Online Shopping

MindfulCart: AI Support for Cognitive Shopping

Chung, Nicholas. [a], *Jin, Jennifer.* [b], *Koo, Seunghyun.* [c] *and Kim, Mira.* [a]
[a] *Department of Computer Science, California State University Fullerton*
E-mail: minsuhk@csu.fullerton.edu, mira.kim@fullerton.edu
[b] *School of Computer Science and Engineering, California State University San Bernardino*
E-mail: jennifer.jin@csusb.edu
[c] *School of AI Convergence, Sungshin Women's University*
E-mail: 20231313@sungshin.ac.kr

Abstract

Online shopping requires sustained attention, working memory, constraint tracking, and comparative reasoning. These types of cognitive functions are often compromised in older adults experiencing mild cognitive impairment (MCI). Despite the rapid growth of e-commerce, cognitive accessibility remains underdeveloped relative to support for visual and motor impairments. This paper presents the design, implementation, and usability evaluation of MindfulCart, an AI-assisted shopping interface that provides structured cognitive scaffolds to support memory, comparison, and decision confidence. A formative usability study evaluated usability, cognitive load, decision confidence, and task success. Preliminary findings suggest reductions in perceived mental effort and confusion and increased decision confidence. Future work will expand the participant pool and explore adaptive personalization mechanisms.

Keywords: Cognitive Scaffolds, MCI, Cognitive Impairment, Online Shopping

Introduction

Online shopping has become a central component of daily life, particularly for older adults who increasingly rely on digital platforms for household goods, medical supplies, and health-related products. While e-commerce environments are designed for convenience and scale, they impose substantial cognitive demands. Users must maintain task goals (e.g., budget constraints, must-haves), compare multi-attribute specifications, track previously viewed alternatives, filter irrelevant information, and revise decisions as new information and/or products emerge (Carpenter and Yoon 2011; Hettich et al. 2018; Acosta-Vargas et al. 2022).

Mild Cognitive Impairment (MCI) affects approximately 10 to 20% of adults over age 65 and is characterized by measurable decline in memory and/or executive functioning that exceeds typical aging while not meeting criteria for dementia. Although individuals with MCI often remain independent in daily living, they may struggle with cognitively complex instrumental activities such as financial management, medication tracking, and consumer decision-making (Petersen 2011; Langa and Levine 2014; Manly et al. 2022). Shopping, particularly online shopping, represents a cognitively dense activity requiring simultaneous constraint tracking and uncertainty management. Errors in such contexts can lead to financial strain, frustration, reduced autonomy, or withdrawal from digital platforms altogether.

Despite the rapid growth of digital commerce, accessibility research has disproportionately emphasized visual and motor impairments (e.g., screen reader compatibility, contrast ratios, keyboard navigation). Many modern e-commerce sites increase cognitive burden through:

- Expansive filtering systems with dozens of parameters
- Large specification tables requiring cross-tab comparison
- Dynamic content updates and recommendation carousels
- Frequent navigation loops between product detail pages

These design patterns presume intact executive functioning and impose substantial self-regulation demands. Users must internally maintain constraints (e.g., "under $120 and quiet"), remember which items were previously viewed, and mentally compare specifications across multiple browser tabs. Such demands may disproportionately disadvantage older adults experiencing subtle cognitive decline.

Research in ecologically valid cognitive assessment demonstrates that shopping tasks are sensitive indicators of early cognitive impairment. Virtual supermarket paradigms show that planning, list management, and constraint adherence degrade in individuals with MCI, even when traditional memory tests show only mild deficits. These findings suggest that everyday consumer decision-making provides a realistic and meaningful context in which cognitive supports may be beneficial (Yan et al. 2021; Zygouris et al. 2022; Rydzewska et al. 2024).

Parallel work in Human-Computer Interaction (HCI) demonstrates that external memory aids, such as digital reminders and structured note systems, can enhance independence and reduce stress in individuals with cognitive limitations (Lanzi and Bourgeois 2020; Mason et al. 2012; Holthe et al. 2022; Wilson et al. 2022). However, these supports are typically deployed as standalone tools rather than integrated directly into decision environments. As a result, users must coordinate across systems, introducing additional cognitive load. Recent advances in artificial intelligence (AI), particularly large language models (LLMs), provide new opportunities to embed dynamic cognitive scaffolds directly into interactive systems. Unlike traditional recommender systems that prioritize predictive personalization (e.g., "customers also bought"), AI-driven scaffolds can provide:

- Goal tracking ('You selected quiet and under $120')
- Constraint monitoring ('This item exceeds your budget')
- Bounded comparison ('Here are the 3 key differences')
- Plain-language summaries of tradeoffs
- Explainable reasoning tied to user-defined preferences

This paper introduces **MindfulCart**, an AI-assisted e-commerce interface designed specifically to reduce cognitive friction during product comparison and selection. The system integrates six scaffolding components: goal locking, orientation tracking, bounded comparison, memory offloading via shortlists, explainable summaries, and user-controlled notes. The interface was implemented within a prototype environment featuring a simplified air purifier store. The research addresses three primary questions:

1. Can AI-based cognitive scaffolds reduce perceived mental effort during online shopping tasks?
2. Do structured supports improve constraint adherence and reduce navigation churn?
3. Does explainable assistance increase decision confidence without undermining autonomy?

Related Work

MindfulCart sits at the intersection of (1) ecologically valid shopping-task research in cognitive impairment, (2) cognitive web accessibility and e-commerce barriers, (3) assistive cognition and external memory scaffolds, and (4) human-centered AI and explainability for older adults. This section synthesizes prior work and positions MindfulCart's contribution as embedded, autonomy-preserving cognitive scaffolding inside a commerce decision workflow rather than a conventional recommender (Petersen 2011; Yan et al. 2021; Acosta-Vargas et al. 2022; Lanzi and Bourgeois 2020; Miller 2019; van Leersum 2026).

Shopping Tasks and Cognitive Decline

Instrumental activities such as planning, prospective memory, and multi-step decision-making often decline earlier than basic self-care in aging populations. Shopping tasks are particularly sensitive because they require constraint tracking, working memory maintenance, and tradeoff reasoning. Virtual supermarket paradigms have demonstrated the ability to differentiate older adults with mild cognitive impairment (MCI) from healthy controls, validating shopping environments as cognitively meaningful contexts. However, prior work primarily treats shopping tasks as diagnostic tools rather than assistive systems. MindfulCart reframes the paradigm from assessment to support by integrating cognitive scaffolding into the browsing–comparison–decision process (Petersen 2011; Langa and Levine 2014; Yan et al. 2021; Zygouris et al. 2022; Rydzewska et al. 2024).

Cognitive Accessibility: Standards, Guidance, and Persistent Gaps

Although standards such as WCAG formally include cognitive disabilities, implementation remains inconsistent. Large-scale accessibility audits reveal persistent barriers in commercial websites, including major e-commerce platforms. While compliance discussions often emphasize visual or motor impairments, cognitive accessibility, particularly support for memory load and decision clarity, remains underdeveloped.

Guidelines advocate reducing cognitive load, yet they rarely specify interaction-level supports for complex consumer decisions such as multi-attribute comparison or constraint revision. MindfulCart operationalizes cognitive accessibility through explicit mechanisms: goal locking, persistent context panels, bounded comparisons, and structured summaries (W3C 2024; W3C 2025; Acosta-Vargas et al. 2022; Moreno et al. 2024).

Design Patterns for Cognitive Accessibility and Content Simplification

A growing research thread translates cognitive accessibility principles into reusable design patterns. There has been research on cognitive accessibility design patterns and apply them within a system aimed at making web interaction easier for people with cognitive impairments, including older adults (Moreno et al. 2024). These pattern-oriented contributions reinforce design moves central to MindfulCart: consistency, reduced complexity, simplified presentation, and structured task flow.

MindfulCart extends pattern-based guidance in two ways. First, it applies pattern logic to commerce-specific workflows (compare–decide), where the core burden is not just reading content but maintaining and reconciling constraints across alternatives. Second, it couples these patterns with AI-generated plain-language explanations and 'why this fits' statements to reduce inferential burden during tradeoff evaluation (Moreno et al. 2024; Miller 2019; Barredo Arrieta et al. 2020).

Explainable AI, Trust, and Older Adults' Preferences for Explanations

Even when AI assistance can reduce workload, adoption depends on trust, perceived control, and understanding. Work on explainable AI in aging contexts argues that opaque AI decisions can undermine willingness to use supportive technologies, and that explanation content should match older adults' informational needs and mental models. A recent paper in *The Gerontologist* discusses how older adults think about what AI should explain and why explainability influences trust and willingness to embed AI into daily life. Complementary work examines older adults' preferences for explanations in conversational AI contexts, including how explanation types may affect acceptance, perceived usefulness, and comfort (Miller 2019; Barredo Arrieta et al. 2020; van Leersum 2026; Wong et al. 2025).

MindfulCart's explainability choices, brief 'why this fits' tied to user-locked goals, bounded summaries, and differences-focused comparisons, are designed to support trust without over-automation. The assistant explains tradeoffs and constraint alignment rather than asserting an authoritative best choice.

Much XAI research studies explanations for model decisions in isolation. Less work connects explainability to multi-step decision workflows where the user's goals evolve (e.g., 'change of mind' scenario) and where the system must preserve agency while reducing cognitive friction. MindfulCart evaluates explainability as part of the full browse-compare-decide loop.

Summary and Positioning

Across these literatures, several themes re-occur:

1. Shopping-like tasks are cognitively sensitive for aging/MCI populations.
2. Cognitive accessibility remains under-addressed in practice, including in e-commerce.
3. Design patterns exist for cognitive accessibility, but commerce decision workflows need more explicit support for constraint tracking and bounded comparison.
4. Explainability and trust are essential for older adults' acceptance of AI support.

MindfulCart's contribution is to unify these strands into an implemented e-commerce prototype that provides embedded cognitive scaffolds (goal locking, orientation cues, shortlist memory offload, bounded comparison, and explainable summaries) and evaluates them in realistic shopping tasks.

System Design

MindfulCart was designed as a cognitively supportive e-commerce interface grounded in principles from cognitive aging research, Cognitive Load Theory (Sweller, 1988), distributed cognition, and cognitive accessibility guidelines. Rather than optimizing for engagement, click-through rates, or predictive ranking, the system was explicitly engineered to reduce cognitive friction during multi-attribute consumer decision-making. We define cognitive friction as the mental effort required to

maintain constraints, compare alternatives, manage task state, and recover from disorientation during complex digital tasks. Four design principles guided the system architecture:

1. Externalize Working Memory - Reduce internal memory burden by making constraints, viewed items, and shortlist candidates persistently visible.
2. Bound Cognitive Complexity - Limit simultaneous comparison to three items and restrict comparison attributes to five key variables to prevent overload.
3. Preserve User Agency - Avoid automated decision-making. The system supports but does not override user choice.
4. Provide Explainable Alignment - Tie system suggestions directly to user-defined goals to maintain transparency and trust.

These principles distinguish MindfulCart from typical recommender systems that prioritize predictive ranking rather than cognitive scaffolding.

Implementation Details

MindfulCart was developed using Next.js, with JavaScript and TypeScript serving as the primary languages for both the frontend and backend.

Figure 1. illustrates the main interface of the MindfulCart prototype. The frontend was implemented using Next.js App Router. React was used to build the interactive user interface, and React Hooks were extensively employed to manage application state, behavioral monitoring logic, intervention pop-ups, and the AI assistant panel. TypeScript was incorporated to improve maintainability, enhance type safety, and reduce runtime errors as the interface logic became more complex. For styling, Tailwind CSS and PostCSS were used to maintain visual consistency across pages and reusable components.

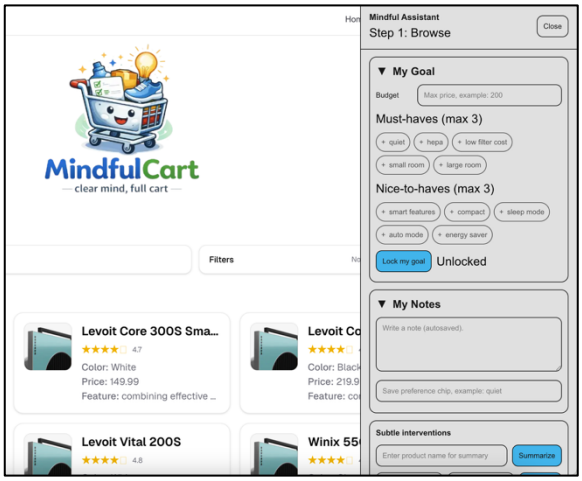


Figure 1. Main Interface of MindfulCart

On the backend, Next.js API routes were used to keep application logic within the same project structure. These routes supported AI assistant functions such as product summarization, goal-alignment checks, and product comparison responses.

Sample product data were stored locally in a CSV file. PapaParse was used to parse the CSV data and convert them into JavaScript objects that could be consumed by the application. User-specific information, including goals, pinned products, comparison selections, and behavioral tracking data, was stored in localStorage.

Prototype Architecture

The prototype was implemented as a controlled web-based e-commerce environment with the following components:

- 41 air purifier products
- Structured dataset - (product, amazon_search_url, price_usd, color, size, data_source, specs, easy_to_install_setup, odor_removal_efficiency, star_review_out_of_5, short_summary)
- Simplified browsing interface
- AI assistant side panel

The air purifier domain was selected for three reasons: (1) Attributes are quantifiable and comparable, (2) Tradeoffs are realistic (noise vs. coverage vs. cost vs. maintenance), and (3) The domain is relevant to older adults (health, indoor air quality). The domain supports systematic evaluation of constraint adherence and tradeoff reasoning under cognitive load.

Data Model and Product Representation

Each product in the system was represented as a structured object containing:

- Unique ID
- Price (continuous)
- Noise level (dB)
- Room coverage (square footage)
- Filter replacement cost
- User rating (1–5 scale)

All attributes were normalized to support weighted scoring and constraint evaluation. For example:

- Price and filter cost were normalized relative to user-specified budget.
- Noise level was inverted when users selected “quiet” as a must-have.

- Room coverage was evaluated against intended room size requirements.

This structured representation enabled deterministic constraint checking and consistent comparison logic independent of AI-generated language output.

Composite Scoring Function

To support transparent and reproducible ranking, MindfulCart used a rule-based composite scoring function rather than a learned ranking model. For each product p, the ranking score was defined as:

$$\text{Score}(p)=0.40M(p)+0.20N(p)+0.25B(p)+0.15S(p)$$

where $M(p)$, $N(p)$, $B(p)$ and $S(p)$ represent must-have match, nice-to-have match, budget proximity, and maintenance simplicity, respectively. Must-have satisfaction received the highest weight to enforce core user constraints. Budget proximity rewarded products that remained within budget and near the target price point, while maintenance simplicity favored products with lower normalized filter replacement cost. This formulation was intentionally lightweight and interpretable, allowing the contribution of each factor to remain explicit.

AI Assistant Modules

The AI assistant functions as a modular cognitive scaffold composed of five core components.

Goal Locking Module

Users specify three components of: (1) Budget, (2) Up to 3 must-haves, and (3) Up to 3 nice-to-haves. Once 'locked', these constraints are persistently displayed, as shown in Figure 2. The system flags violations but does not prevent selection, preserving autonomy. The cognitive rationale was Goal locking reduces constraint drift, where users gradually forget initial priorities as they encounter new information.



Figure 2. Goal Locking Module of MindfulCart

Orientation Tracker

A visible step indicator of (1) Browse → (2) Compare → (3) Decide is implemented. The tracker displays: (1) Number of pinned items and (2) Number of comparisons performed and (3) Current Stage. This module addresses disorientation common in high-choice environments, where users repeatedly reopen pages or lose track of progress. By visualizing task state, the system reduces navigation churn and looping behaviors.

Goal Review and Memory Support via AI Assistant

The AI assistant reduces memory burden by keeping the user's goals continuously visible and easy to revise. Users can quickly review their budget, must-have preferences, nice-to-have preferences, and goal status directly within the assistant panel, reducing the need to remember previously entered criteria while browsing and comparing products. If preferences change, goals can be updated through the same interface without disrupting the shopping flow.

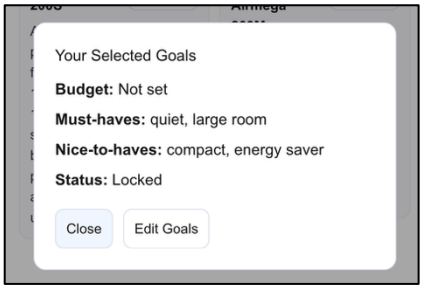


Figure 3. AI-Assisted Goal Review

The assistant also supports bounded decision-making through a shortlist of up to three pinned items. Each item includes the price, three key specifications, and a brief explanation of how it aligns with the user's goals. This design supports recognition over recall, reduces repeated page revisits, and encourages focused comparison without overwhelming the user.

Bounded Comparison Interface

When 2-3 items are selected, we only display key attributes, differences-only comparison, and Plain-language summary generated by LLM ('Model A is quieter but costs more upfront. Model B has cheaper filters but is louder.'). Bounding comparison scope directly addresses intrinsic cognitive load as limiting visible attributes helps prevent extraneous load. The differences-only view reduces visual clutter, and the summary reduces inferential reasoning effort required to interpret numeric tables. Our approach for bounded comparison interface is the following:

- Attribute filtering and scoring are rule-based.
- Natural language summaries are generated via LLM prompt templates.
- Prompts include explicit user priorities to ensure alignment.

Explainable Outputs

Instead of ranking items abstractly, the assistant generates short statements linked to the user's goals. Explainability serves two purposes: (1) Reduce inference burden (2) Increase confidence and transparency, and (3) Prevent over-reliance on opaque algorithmic ranking. These explanations are limited to 1-2 sentences and explicitly reference user-locked constraints. They are generated after scoring based on weighted criteria. The explanation layer translates numeric scoring into interpretable reasoning. We used the following Scoring weights after fine-tuning:

- Must-haves match: 40%
- Nice-to-haves match: 20%
- Budget closeness: 25%
- Low-maintenance simplicity: 15%

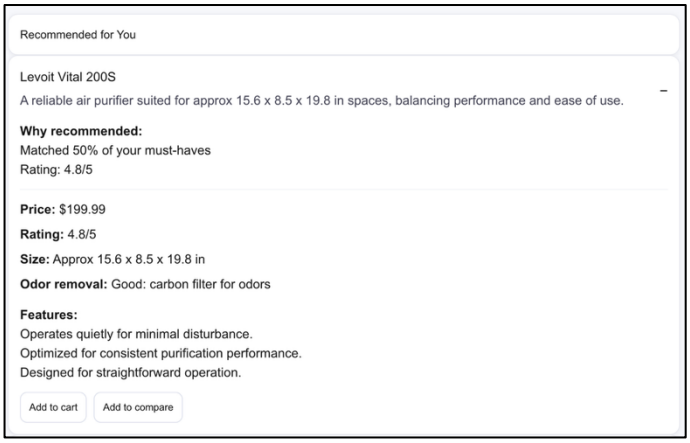


Figure 4. Explainable Outputs of MindfulCart

Interaction Logging Infrastructure

To support objective evaluation, the system logged:

- Page views and navigation transitions
- Time spent per page
- Goal edits, locks, and unlocks
- Must-haves and nice-to-have
- Product pinning and unpinning
- Constraint checks
- Assistant panel opening and closing
- LLM usage
- Notes

Time on task was computed as the interval between task instruction presentation and final confirmation click. Page revisits were defined as returning to a product detail page previously accessed during the same task. This logging architecture enabled behavioral quantification of navigation churn and information retrieval patterns, both of which serve as proxies for working memory strain.

Formative Usability Evaluation

Study Design

To obtain early feedback on the usability and perceived usefulness of MindfulCart, we conducted a small-scale formative usability study with four older adults. The purpose of this evaluation was to examine whether the system's cognitively supportive features were understandable, usable, and potentially helpful during online shopping tasks. Rather than serving as a comprehensive effectiveness study, this pilot was intended to provide initial evidence of feasibility, identify early usability patterns, and inform future refinement.

A within-subject structure was used so that each participant could interact with both interfaces: (1) a baseline shopping interface with standard browsing and comparison functionality, and (2) the MindfulCart interface, which incorporated goal locking, shortlist memory support, bounded comparison, orientation support, and explainable AI summaries. This design allowed participants to directly experience the differences between the two interfaces while reducing variability related to digital literacy, shopping habits, and general comfort with technology.

To reduce simple order effects, the presentation order of the two conditions was alternated across participants. Each session lasted approximately 30 minutes and was conducted individually in a quiet environment. Given the exploratory nature of the study and the limited sample size, the evaluation emphasized usability observations and participant feedback rather than confirmatory statistical analysis.

Participants

Four older adults between the ages of 60 and 71 years participated in the study. Participants were recruited through community networks and word-of-mouth referrals. All reported regular internet use and prior experience with online shopping, which helped ensure that the study primarily captured reactions to the interface design rather than basic unfamiliarity with digital shopping environments.

Participants were informally screened to ensure they did not have severe visual or motor limitations that would substantially interfere with computer use. No formal cognitive or neuropsychological assessments were administered. The goal of the study was not to diagnose cognitive impairment, but rather to gather early feedback from older adult users who represent the intended population for

cognitively supportive shopping tools. Because of the small sample size, the findings are best interpreted as preliminary and formative, offering insight into usability and design feasibility rather than generalizable evidence of effectiveness.

Task Design

The study tasks were designed to align closely with the core interaction flow of MindfulCart: Browse → Compare → Decide. In the prototype, users define and lock their shopping goals, browse a structured set of air purifier products, pin up to three candidate items, compare shortlisted products using a bounded comparison interface, and make a final decision with support from brief AI-generated explanations linked to their goals.

Participants completed three sequential tasks in each condition: (Task 1) Specify and lock a shopping goal by entering a budget, must-haves and nice-to-haves, (Task 2) Browse the product set and pin up to three items that match their criteria. This task focused on bounded comparison and tradeoff reasoning. (Task 3) Given a revised preference scenario, confirm/revise their earlier choice. This final task examined decision revision and the extent to which the interface supported users in adapting to changing priorities while remaining oriented to their original goals. Together, these tasks were intended to reflect the actual affordances of the MindfulCart system and to evaluate whether its cognitive scaffolding features supported older adults during realistic online shopping decisions.

Usability Evaluation and Feedback

Because this study was designed as a formative usability evaluation, the analysis focused on descriptive usability observations, interaction patterns, and brief participant feedback rather than formal statistical testing. The purpose of the evaluation was to examine whether MindfulCart's cognitively supportive features were understandable, usable, and helpful for older adults during multi-attribute shopping tasks.

The evaluation considered several broad indicators of usability and decision support. These included whether participants were able to complete the shopping tasks, how smoothly they moved through the Browse-Compare-Decide workflow, the extent to which they revisited previously viewed products, whether they used the shortlist and comparison features as intended, and whether their final selections remained aligned with their stated goals.

In addition to these observational measures, brief post-task feedback was collected to understand participants' subjective experience with the system. Participants were asked about ease of use, clarity of the workflow, helpfulness of the shortlist and comparison view, usefulness of the AI-generated explanations, confidence in their choice, and perceived mental effort.

Preliminary Observations

Overall, participant feedback suggested that the structured workflow of MindfulCart was understandable and that several of its cognitive scaffolding features were useful during shopping. In particular, the persistent display of goals appeared to help participants stay focused on their priorities, while the shortlist reduced the need to repeatedly remember or relocate previously viewed products. The bounded comparison view also appeared to make product tradeoffs easier to review without becoming overly complex. Participants generally responded positively to the AI-generated explanations when they clearly linked product recommendations to their stated goals.

Observationally, participants appeared to benefit from having a more structured decision process than that offered by a conventional browsing interface alone. Features such as goal locking, shortlist maintenance, and bounded comparison seemed to support a more organized shopping experience by reducing the need to mentally track multiple products and criteria at once. Although these patterns are preliminary, they suggest that cognitively supportive design features may help older adults navigate multi-attribute shopping tasks with greater clarity and confidence.

Conclusion and Future Work

MindfulCart demonstrates that structured AI scaffolding can reduce measurable cognitive friction in multi-attribute online shopping tasks among older adults. Across participants, the AI-supported interface improved constraint adherence, reduced navigation backtracking, lowered perceived mental effort, and increased decision confidence. These improvements were achieved without automating final choice, thereby preserving autonomy and transparency.

From a technical standpoint, the system illustrates the feasibility of a hybrid architecture combining deterministic scoring with constrained natural language generation. This separation ensures

explainability while leveraging AI to reduce interpretive burden. Unlike predictive recommender systems optimized for engagement, MindfulCart prioritizes cognitive stability and goal alignment.

Despite promising results, several limitations warrant consideration. The sample size was small and not clinically diagnosed for mild cognitive impairment. The study occurred in a controlled prototype environment without real financial consequences, and the domain was limited to air purifiers. Furthermore, longitudinal effects on digital independence were not assessed.

Future work will extend this research along three dimensions. First, a larger-scale study with over 50 participants including clinically confirmed MCI participants will allow more precise estimation of effect sizes and subgroup analyses. Second, adaptive scaffolding mechanisms will be explored, in which the system dynamically adjusts the level of cognitive support based on observed behavioral indicators such as excessive revisits or prolonged decision latency. Third, the framework will be generalized to additional domains, including financial products, and health-related decision contexts.

Longer-term research will examine ethical considerations of AI memory augmentation, particularly how persistent preference tracking and automated reminders can support independence without creating over-reliance. The ultimate goal is to develop cognitively adaptive digital environments that enhance autonomy, reduce decision fatigue, and promote inclusive access to essential online services.

References

Acosta-Vargas, P., Salvador-Acosta, B., Salvador-Ullauri, L., and Jadán-Guerrero, J. 2022. "Accessibility Challenges of E-Commerce Websites," *PeerJ Computer Science* (8), p. e891.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. 2020. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI," *Information Fusion* (58), pp. 82–115.

Carpenter, S. M., and Yoon, C. 2011. "Aging and Consumer Decision Making," *Annals of the New York Academy of Sciences* (1235:1), pp. E1–E12.

Hettich, D., Hattula, S., and Bornemann, T. 2018. "Consumer Decision-Making of Older People: A 45-Year Review," *The Gerontologist* (58:6), pp. e349–e368.

Holthe, T., Halvorsrud, L., and Lund, A. 2022. "Digital Assistive Technology to Support Everyday Living in Community-Dwelling Older Adults with Mild Cognitive Impairment and Dementia," *Clinical Interventions in Aging* (17), pp. 519–544.

Kim, J., Maathuis, H., and Sent, D. 2024. "Human-Centered Evaluation of Explainable AI Applications: A Systematic Review," *Frontiers in Artificial Intelligence* (7), p. 1456486.

Langa, K. M., and Levine, D. A. 2014. "The Diagnosis and Management of Mild Cognitive Impairment: A Clinical Review," *JAMA* (312:23), pp. 2551–2561.

Lanzi, A., and Bourgeois, M. S. 2020. "Structured External Memory Aid Treatment for Mild Cognitive Impairment," *American Journal of Speech-Language Pathology* (29:1S), pp. 474–484.

Manly, J. J., Jones, R. N., Langa, K. M., Ryan, L. H., Levine, D. A., McCammon, R., Heeringa, S. G., and Weir, D. 2022. "Estimating the Prevalence of Dementia and Mild Cognitive Impairment in the US: The 2016 Health and Retirement Study Harmonized Cognitive Assessment Protocol Project," *JAMA Neurology* (79:12), pp. 1242–1249.

Mason, S., Craig, D., O'Neill, S., Donnelly, M., and Nugent, C. 2012. "Electronic Reminding Technology for Cognitive Impairment," *British Journal of Nursing* (21:14), pp. 855–861.

Miller, T. 2019. "Explanation in Artificial Intelligence: Insights from the Social Sciences," *Artificial Intelligence* (267), pp. 1–38.

Moreno, L., Petrie, H., Martínez, P., and Alarcon, R. 2024. "Designing User Interfaces for Content Simplification Aimed at People with Cognitive Impairments," *Universal Access in the Information Society* (23:1), pp. 99–117.

Petersen, R. C. 2011. "Clinical Practice. Mild Cognitive Impairment," *The New England Journal of Medicine* (364:23), pp. 2227–2234.

Rotondi, A. J., Sinkule, J., Haas, G. L., Spring, M. B., Litschge, C. M., Newhill, C. E., Ganguli, R., and Anderson, C. M. 2007. "Designing Websites for Persons with Cognitive Deficits: Design and Usability of a Psychoeducational Intervention for Persons with Severe Mental Illness," *Psychological Services* (4:3), pp. 202–224.

Rydzewska, K., Nielek, R., Pawłowska, J., Wierzbicki, A., and Sedek, G. 2024. "Performance of Aging Consumers in an E-Commerce Product Choice Task: The Role of Working Memory and Decision Strategies," *PLOS ONE* (19:6), p. e0303806.

van Leersum, C. M. 2026. “What Should AI Explain to You? Visions of Older Adults Regarding Explainable AI,” *The Gerontologist* (66:2), p. gnaf308.

Wilson, S. A., Byrne, P., Rodgers, S. E., and Maden, M. 2022. “A Systematic Review of Smartphone and Tablet Use by Older Adults With and Without Cognitive Impairment,” *Innovation in Aging* (6:2), p. igac002.

World Wide Web Consortium (W3C). 2024. “Web Content Accessibility Guidelines (WCAG) 2.2,” W3C Recommendation, December 12.

Wong, A. K. C., Lee, J. H. T., Zhao, Y., Lu, Q., Yang, S., and Hui, V. C. C. 2025. “Exploring Older Adults’ Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study,” *JMIR Aging* (8), p. e66778.

Yan, M., Yin, H., Meng, Q., Wang, S., Ding, Y., Li, G., Wang, C., and Chen, L. 2021. “A Virtual Supermarket Program for the Screening of Mild Cognitive Impairment in Older Adults: Diagnostic Accuracy Study,” *JMIR Serious Games* (9:4), p. e30919.

Zygouris, S., Segkouli, S., Triantafyllidis, A., Giakoumis, D., Tsolaki, M., Votis, K., and Tzovaras, D. 2022. “Usability of the Virtual Supermarket Test for Older Adults with and without Cognitive Impairment,” *Journal of Alzheimer’s Disease Reports* (6:1), pp. 229–234.

B3

Scalable Sentiment Analysis for AI-Powered E-Commerce: Comparing ML and BERT-Based NLP Models on Large-Scale Amazon Book Reviews

Hsiu-Ping Lin (California State University Los Angeles), Sunnyoung Lee (Dongguk University) and Jongwook Woo (California State University Los Angeles).

The rapid expansion of AI-driven e-commerce has intensified the need for accurate sentiment analysis of customer reviews to enhance product recommendations and personalized shopping experiences. This study compares traditional machine learning (ML) models and fine-tuned BERT for binary sentiment classification using the Amazon Books Reviews dataset of approximately 3 million reviews. We employ PySpark on an AWS EMR cluster to build four traditional models – Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron Neural Network (MLP-NN) – and fine-tune BERT on sample datasets of 10,000 and 50,000 reviews via Google Colab. Models are evaluated using AUC and Precision. Results show that fine-tuned BERT consistently outperforms all traditional models, with performance improving as sample size increases, while LR achieves the best computational efficiency. Training on full Review Text substantially outperforms Review Summary across all models. These findings offer practical guidance for e-commerce platforms deploying AI-powered recommendation systems.

KEY WORDS

AI-Driven E-Commerce, Sentiment Analysis, Customer Experience, BERT, Machine Learning, NLP, Big Data

Scalable Sentiment Analysis for AI-Powered E-Commerce: Comparing ML and BERT-Based NLP Models on Large-Scale Amazon Book Reviews

Lin, H. P.[a], Lee, S. N.[b], and Woo, J. W.[a]
[a] Department of Information Systems, California State University Los Angeles
E-mail: hsiuping01@gmail.com, jwoo5@calstatela.edu
[b] Department of Business Administration, Dongguk University
E-mail: enhasu19@dongguk.edu

Abstract

The rapid expansion of AI-driven e-commerce has intensified the need for accurate sentiment analysis of customer reviews to enhance product recommendations and personalized shopping experiences. This study compares traditional machine learning (ML) models and fine-tuned BERT for binary sentiment classification using the Amazon Books Reviews dataset of approximately 3 million reviews. We employ PySpark on an AWS EMR cluster to build four traditional models — Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron Neural Network (MLP-NN) — and fine-tune BERT on sample datasets of 10,000 and 50,000 reviews via Google Colab. Models are evaluated using AUC and Precision. Results show that fine-tuned BERT consistently outperforms all traditional models, with performance improving as sample size increases, while LR achieves the best computational efficiency. Training on full Review Text substantially outperforms Review Summary across all models. These findings offer practical guidance for e-commerce platforms deploying AI-powered recommendation systems.

Keywords: AI-Driven E-Commerce, Sentiment Analysis, Customer Experience, BERT, Machine Learning, NLP, Big Data

Introduction

The rapid growth of AI-driven e-commerce has fundamentally transformed how consumers discover and evaluate products. Among the various signals that shape purchasing decisions, customer reviews play a particularly influential role. On platforms such as Amazon, millions of users rely on review content to assess product quality, making accurate sentiment classification a strategic priority for recommendation systems and personalized shopping experiences (Xing and Zhan 2015). However, the sheer volume and unstructured nature of review data pose significant analytical challenges that demand both sophisticated natural language processing (NLP) techniques and scalable computing infrastructure.

Traditional machine learning (ML) models, such as Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron Neural Network (MLP-NN), have long been applied to text classification tasks in e-commerce settings (Mishra et al. 2019; Mishra et al. 2021). These models are computationally efficient and interpretable, but they rely on bag-of-words or TF-IDF representations that inherently fail to capture contextual nuances in language. Book reviews, in particular, frequently exhibit sentiment transitions — a review may open with praise before closing with critical feedback — making context-sensitive understanding essential for accurate classification (Wang and Shin 2019). Recent benchmarking studies have further confirmed that while traditional ML models remain

competitive in terms of speed and scalability, they consistently underperform transformer-based models on nuanced sentiment tasks in e-commerce contexts (Sharaf et al. 2024; Erdogan et al. 2025).

The emergence of transformer-based Large Language Models (LLMs), most notably BERT (Bidirectional Encoder Representations from Transformers), has addressed many of these limitations. By pre-training on large corpora and capturing bidirectional contextual relationships between words, BERT and its variants have demonstrated state-of-the-art performance across a wide range of NLP tasks (Chen et al. 2023). Fine-tuned BERT models have shown particular strength in e-commerce review analysis, where contextual understanding of nuanced language directly translates to higher classification accuracy (Jahan et al. 2025). The performance advantage of fine-tuned BERT over both traditional ML models and zero-shot LLMs has been consistently documented across multiple e-commerce domains, while its computational efficiency relative to larger generative models makes it a practical choice for production deployment (Roumeliotis et al. 2024; Yilmaz and Kose 2025).

A further challenge lies in the scale of modern e-commerce review datasets. The Amazon Books Reviews dataset, for instance, contains approximately 3 million records totaling 2.7 GB, far exceeding the processing capacity of conventional single-node systems. Big Data frameworks such as Apache Spark, deployed on distributed cloud platforms like AWS Elastic MapReduce (EMR), offer a scalable solution for processing such datasets efficiently (Mishra and Woo 2021). Integrating these platforms into the NLP pipeline is therefore not merely a technical convenience but a practical necessity for production-grade e-commerce AI systems (Yadav et al. 2021).

Against this backdrop, this study addresses the following research questions: (1) How do traditional ML models compare to fine-tuned BERT in predicting customer sentiment from Amazon book reviews? (2) How does model performance change as the fine-tuning sample size scales from 10,000 to 50,000 reviews? (3) Does using full Review Text yield meaningfully better classification accuracy than using Review Summary across both model families?

To answer these questions, we construct binary sentiment classification models using LR, DT, RF, and MLP-NN on the full 3-million-record dataset via PySpark on AWS EMR, and fine-tune BERT on sample datasets of 10,000 and 50,000 reviews using Google Colab with a T4 GPU. Model performance is evaluated using AUC and Precision, with particular emphasis on minimizing false positives — misclassifying low-rated books as positive recommendations — given the direct impact of this error type on customer satisfaction (Mishra et al. 2020).

The remainder of this paper is organized as follows. Section 2 reviews related work on sentiment analysis and NLP model comparisons in e-commerce contexts. Section 3 describes the dataset, system architecture, and data preprocessing procedures. Section 4 presents the experimental results across all models and conditions. Section 5 discusses the findings and their implications for AI-driven e-commerce platforms, followed by conclusions and directions for future research.

Related Work

Research on e-commerce sentiment analysis has progressed from traditional ML classifiers to transformer-based LLMs. Xing and Zhan (2015) and Wang and Shin (2019) established early benchmarks using feature engineering approaches, while Mishra et al. (2019) extended this to large-scale Amazon review data, confirming the utility of Logistic Regression in high-volume settings despite its inability to capture contextual language nuances.

Transformer-based BERT models have since demonstrated consistent superiority over traditional ML approaches in e-commerce sentiment tasks. Norinder and Norinder (2022) showed that deep learning improves accuracy on imbalanced review datasets, and Sharaf et al. (2024) confirmed that BERT-based models outperform traditional classifiers while identifying computational cost as a key deployment consideration. Jahan et al. (2025) and Yilmaz and Kose (2025) further validated fine-tuned BERT's advantage across multiple domains and languages. For large-scale processing, Mishra and Woo (2021) demonstrated that PySpark on distributed clusters is essential for multi-gigabyte datasets, and Roumeliotis et al. (2024) showed that fine-tuned BERT achieves lower error rates and faster inference than generative LLMs such as GPT-3.5 and LLaMA-2.

Despite this progress, prior studies rarely combine full-dataset traditional ML modeling with BERT fine-tuning in a unified framework, and the effect of sample size scaling and input feature granularity — Review Summary versus full Review Text — remains underexplored for Amazon Books reviews. The present study directly addresses these gaps.

Data, Architecture, and Preprocessing

This section describes the dataset used in this study, the system architecture employed for large-scale data processing and model training, and the data preprocessing procedures applied to prepare the data for classification.

Dataset

The dataset used in this study is the "Amazon Books Reviews" dataset obtained from Kaggle (Mohamed 2023). The primary file, *Books_rating.csv*, is 2.7 GB in size and contains approximately 3 million book reviews. Each record includes the fields listed in Table 1. The Review Text and Review Summary fields serve as the primary textual features for sentiment classification, while Price and Review Time are used as supplementary numerical features. An additional file, *Books_data.csv*, was available but not utilized in this study.

Table 1. Dataset Fields in Books_rating.csv

Field	Description
ID	Unique review identifier
Title	Book title
Price	Book price
User ID	Anonymized user identifier
Profile Name	Reviewer profile name
Review Summary	Short summary of the review (used as textual feature)
Review Text	Full review text (used as textual feature)
Review Helpfulness	Helpfulness votes (numerator/denominator)
Rating	User rating from 1 to 5 (converted to binary label)



Figure 1. System Architecture Overview

To construct a binary classification target, ratings of 1–3 are mapped to label 0 (negative/neutral) and ratings of 4–5 are mapped to label 1 (positive). After preprocessing, approximately 21% of records carry label 0 and 79% carry label 1, reflecting the strongly positive skew typical of e-commerce review datasets.

System Architecture

Given the large-scale nature of the dataset, two separate computing environments are employed depending on the model type, as illustrated in Figure 1.

For traditional ML models, PySpark is deployed on an AWS Elastic MapReduce (EMR) cluster to process the full 2.7 GB dataset in a distributed manner. The EMR cluster configuration is detailed in Table 2. PySpark's distributed computing framework allows parallel processing across multiple nodes, significantly reducing computation time and enabling linear scalability as data volume grows.

Table 2. AWS EMR Cluster Specification

Specification	Details
Number of Nodes	5
Hadoop Cluster Version	3.1.2
Spark Version	3.0.2
CPU per Node	2.4 GHz, 64 cores

For BERT fine-tuning, PySpark does not natively support large language model training. Therefore, sample datasets of 10,000 and 50,000 records are extracted from the full dataset and fine-tuned using a T4 GPU on Google Colab. The T4 GPU provides the parallel processing capability necessary for efficient transformer model training, enabling faster convergence during fine-tuning compared to CPU-based configurations.

Data Preprocessing

Preprocessing procedures differ between the traditional ML pipeline on PySpark and the BERT fine-tuning pipeline on Google Colab.

Traditional ML Preprocessing (PySpark). Three features are used to train the traditional ML models: two numerical features (Price and Review Time) and one textual feature (either Review Summary or Review Text). For numerical features, missing values are replaced with the respective column mean, after which MinMaxScaler normalization is applied. For the textual feature, records with missing values are removed, and the following NLP preprocessing steps are applied sequentially: Tokenizer, StopWordsRemover, CountVectorizer, and Inverse Document Frequency (IDF) weighting. The vocabulary size (vocabSize) in CountVectorizer is set to 2^{16} (65,536) by default, though this value is reduced to 4,096 for the MLP-NN model due to memory constraints.

BERT Fine-Tuning Preprocessing (Google Colab). For BERT fine-tuning, only the textual feature (Review Summary or Review Text) is used as input. Text is tokenized using the BERT tokenizer from Hugging Face, with a maximum sequence length applied to manage GPU memory. Fine-tuning is conducted with the following hyperparameter settings: learning rate of 2e-5, training batch size of 16, evaluation batch size of 16, 2 training epochs, and weight decay of 0.01.

Experimental Results

This section presents the classification performance of all models across two experimental conditions: traditional ML models trained on the full dataset using PySpark on AWS EMR, and fine-tuned BERT models trained on sample datasets of 10,000 and 50,000 reviews using Google Colab. Performance is reported in terms of AUC and Precision, with computation time (CT) included to assess practical

deployability. Results are organized by feature type — Review Summary versus Review Text — to evaluate the impact of input feature granularity.

Traditional ML Models — Review Summary Feature

Table 3 presents the performance of the four traditional ML models trained using Review Summary as the textual feature, along with Price and Review Time as numerical features. Bold values indicate the best performance per column while *CT* means *Computation Time*.

Table 3. Performance of Traditional ML Models Using Review Summary Feature

Algorithm	Vocab Size	CT (min)	AUC (%)	Precision (%)
LR	65,536	**21**	**86.76**	**87.55**
DT	65,536	1,203	49.20	81.85
RF	65,536	977	74.21	79.75
MLP-NN	4,096	837	83.40	85.61

Among the traditional models, LR achieves the highest AUC (86.76%) and Precision (87.55%) while also requiring the shortest computation time of 21 minutes — making it the most efficient model in this condition. MLP-NN delivers the second-highest AUC (83.40%) and Precision (85.61%), despite operating with a reduced vocabulary size of 4,096 due to memory constraints. DT shows the lowest AUC (49.20%), suggesting it struggles to capture the sentiment nuances present in short review summaries. RF, while computationally expensive at 977 minutes, achieves only moderate AUC (74.21%) and Precision (79.75%).

Traditional ML Models — Review Text Feature

Table 4 presents the results when the richer Review Text feature replaces Review Summary, while keeping Price and Review Time as numerical features.

Table 4. Performance of Traditional ML Models Using Review Text Feature

Algorithm	Vocab Size	CT (min)	AUC (%)	Precision (%)
LR	65,536	**93**	**90.02**	**89.32**
DT	25,000	343	55.92	81.91
RF	65,536	1,044	80.88	79.77
MLP-NN	4,096	823	88.43	89.25

Switching to full Review Text consistently improves model performance across all algorithms. LR again achieves the highest AUC (90.02%) and Precision (89.32%), though at the cost of a 342.86% increase in computation time compared to using Review Summary (from 21 to 93 minutes). MLP-NN shows the second-best Precision (89.25%) and a substantial AUC improvement from 83.40% to 88.43%, demonstrating that richer textual input disproportionately benefits neural models. DT improves its AUC from 49.20% to 55.92% but remains the weakest performer overall. Notably, DT's vocabulary size is reduced to 25,000 in this condition to manage memory constraints, which may partially limit its performance.

BERT Fine-Tuning Results

Table 5 presents the performance of fine-tuned BERT across two sample sizes (10,000 and 50,000) and two feature types (Review Summary and Review Text). Since BERT operates at the token level through the Hugging Face tokenizer, Vocab Size is not applicable in this context.

Table 5. BERT Fine-Tuning Performance by Sample Size and Feature Type

Algorithm	Sample Size	CT (min)	AUC (%)	Precision (%)
BERT	10,000 (Summary)	**2.5**	80.19	79.23
BERT	10,000 (Text)	19.4	88.31	87.42
BERT	50,000 (Summary)	9.0	83.19	82.03
BERT	50,000 (Text)	107.3	**90.93**	**89.81**

Fine-tuned BERT consistently outperforms all traditional ML models in AUC and Precision when trained on Review Text. With 50,000 samples and Review Text, BERT achieves the highest AUC (90.93%) and Precision (89.81%) of all models in this study. Increasing the sample size from 10,000 to 50,000 yields meaningful gains: AUC improves from 88.31% to 90.93% (+2.62%) and Precision from 87.42% to 89.81% (+2.39%) under the Review Text condition. Similarly, switching from Review Summary to Review Text substantially boosts BERT's performance, with AUC increasing by 8.12% and Precision by 8.19% at the 10,000-sample level. These gains are consistently larger than those observed for traditional ML models under the same feature switch, confirming that BERT's transformer-based architecture is better equipped to exploit the richer semantic content in full review texts.

Cross-Model Comparison

Across all experimental conditions, three findings stand out. First, LR is the most competitive traditional ML model, delivering the highest AUC and Precision among traditional approaches with the shortest computation time — making it a strong baseline for large-scale deployment. Second, fine-tuned BERT with Review Text and 50,000 samples achieves the best overall performance (AUC: 90.93%, Precision: 89.81%), surpassing all traditional ML models in predictive accuracy. However, BERT's computation time of 107.3 minutes for this condition represents a significant cost compared to LR's 93 minutes on the full 3-million-record dataset, highlighting an important efficiency trade-off. Third, Review Text consistently outperforms Review Summary across all models and sample sizes, confirming that the depth and completeness of textual input is a critical driver of sentiment classification accuracy in e-commerce review analysis.

Discussion and Conclusion

This study compared traditional ML models and fine-tuned BERT for binary sentiment classification of Amazon Books reviews across two feature types (Review Summary vs. Review Text) and two BERT sample sizes (10,000 and 50,000). Three key findings emerged from the experiments.

Among traditional ML models trained on the full 3-million-record dataset via PySpark on AWS EMR, LR consistently achieved the highest AUC and Precision under both feature conditions while requiring the shortest computation time, confirming it as the most practical baseline for large-scale e-commerce deployment (Mishra et al. 2019; Sharaf et al. 2024). Fine-tuned BERT with 50,000 samples and Review Text achieved the highest overall performance (AUC: 90.93%, Precision: 89.81%), with performance improving as sample size scaled from 10,000 to 50,000 (Jahan et al. 2025). However, the performance gap over LR remains narrow — approximately 0.91% in AUC and 0.49% in Precision — raising practical questions about whether this marginal gain justifies BERT's substantially higher computational cost. Across all models, switching from Review Summary to Review Text consistently improved performance, with BERT benefiting disproportionately due to its bidirectional contextual encoding (Erdogan et al. 2025).

These findings carry direct implications for AI-driven e-commerce platforms. LR on distributed infrastructure is better suited for real-time inference at scale, while BERT is more appropriate for offline batch processing where marginal accuracy gains carry higher business value. Regardless of model

choice, Big Data infrastructure such as PySpark on AWS EMR is a prerequisite for production-scale NLP deployment (Mishra and Woo 2021), and preserving full review text over summary fields is a strategically important design decision as AI-driven recommendation systems become central to e-commerce competitiveness (Roumeliotis et al. 2024).

This study has several limitations. The 79% positive label imbalance warrants future exploration of resampling techniques such as SMOTE, and BERT fine-tuning was constrained to 50,000 samples due to computational limitations, leaving performance scaling at larger sizes unexplored. Future research should extend experiments to larger fine-tuning datasets, incorporate recent generative LLMs such as GPT-4o and LLaMA-3 for updated benchmarking, and validate the framework across additional Amazon product categories to strengthen generalizability.

References

Chen, Y., Juan, T., and Jung, H. 2023. "Text Classification on Social Network Platforms Based on Deep Learning Models," *Journal of Information and Communication Convergence Engineering* (21:1), pp. 9-16.

Erdogan, B., Calik, N., and Kaya, T. 2025. "Transformer and Pre-Transformer Model-Based Sentiment Prediction with Various Embeddings: A Case Study on Amazon Reviews," *Entropy* (27:12), p. 1202.

Jahan, N., Ahamed, J., and Nandi, D. 2025. "Enhancing E-commerce Sentiment Analysis with Advanced BERT Techniques," *International Journal of Information Engineering and Electronic Business* (17:3), pp. 49-61.

Mishra, M., Chopde, J., Shah, M., Parikh, P., Babu, R. C., and Woo, J. 2019. "Big Data Predictive Analysis of Amazon Product Review," in *Proceedings of the 14th Asia Pacific International Conference on Information Science and Technology (APIC-IST)*, pp. 141-147.

Mishra, M., Kang, M., and Woo, J. 2020. "Leveraging Big Data for Spark Deep Learning to Predict Rating," *Journal of Internet Computing and Services* (21:6), pp. 33-39.

Mishra, M., and Woo, J. 2021. "Predicting the Ratings of Amazon Products Using Big Data," *WIREs Data Mining and Knowledge Discovery* (11:3), e1400.

Norinder, U., and Norinder, P. 2022. "Predicting Amazon Customer Reviews with Deep Confidence Using Deep Learning and Conformal Prediction," *Journal of Management Analytics* (9:1), pp. 1-16.

Roumeliotis, K., Tselikas, N., and Nasiopoulos, D. 2024. "LLMs in E-Commerce: A Comparative Analysis of GPT and LLaMA Models in Product Review Evaluation," *Natural Language Processing Journal* (6), 100056.

Sharaf, M., Hemdan, E. E.-D., El-Sayed, A., and El-Bahnasawy, N. A. 2024. "Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers," *Mathematics* (12:15), p. 2403.

Wang, G., and Shin, S. Y. 2019. "An Improved Text Classification Method for Sentiment Classification," *Journal of Information and Communication Convergence Engineering* (17:1), pp. 41-48.

Xing, F., and Zhan, J. 2015. "Sentiment Analysis Using Product Review Data," *Journal of Big Data* (2:5).

Yadav, S., Muralidharan, S., Lee, S., and Woo, J. 2021. "Scalable Predictive Analysis for Airbnb Listing Rating," in *Proceedings of the 13th International Conference on Internet (ICONI)*, pp. 370-372.

Yilmaz, O., and Kose, U. 2025. "AI-Based Sentiment Analysis of E-Commerce Customer Feedback: A Bilingual Parallel Study on the Fast Food Industry in Turkish and English," *Journal of Theoretical and Applied Electronic Commerce Research* (20:4), p. 294.

B3

Multidimensional Rule-based Grocery Product Classification

Ruobing Liu (Inha University) and Hyunwoo Lim (Inha University).

Situated between volatile supplier conditions and highly fragmented retail demand, grocery wholesalers face greater inventory management challenges than individual stores. This study develops a systematic, multidimensional framework to classify wholesale grocery products based on demand characteristics and proposes differentiated inventory control policies for each resulting profile. Using real-world data from a wholesale logistics center in South Korea, the study finds that most perishable SKUs fall into the non-moving (N) category, and that high seasonality is predominantly associated with low inventory turnover. Case-based analysis further demonstrates that differentiated replenishment strategies are necessary across FSN categories.

KEY WORDS

SKU classification, Inventory profiling, Inventory control, FSN, Grocery products

B3

Does Macroeconomic Uncertainty Drive Online Purchasing? Evidence from Daily Coupang Data.

Jae Ik Ahn, Gyoo Gun Lim (Hanyang University)

This study examines whether macroeconomic uncertainty affects e-commerce demand in South Korea using daily transaction data from Coupang, a leading online platform. Rather than focusing on product substitution, this study investigates channel substitution, arguing that consumers may shift toward online purchasing when economic conditions deteriorate. Using a time-series regression framework with daily purchase amount and purchase count as dependent variables, and the News Sentiment Index, KOSPI, and gold prices as macroeconomic indicators, the analysis covers the period from January 2021 to January 2026. Results show that declining economic sentiment and rising gold prices are associated with increased online purchasing activity, supporting the view that e-commerce functions as a defensive consumption channel under uncertainty. These findings extend the lipstick effect literature by shifting the analytical focus from product choice to channel choice.

KEY WORDS

E-commerce, Lipstick effect, Macroeconomic uncertainty, Channel substitution, Consumer sentiment

B3

Can Trust Signals Defend Against Negative Reviews? Evidence from the South Korean E-Commerce Market.

Jae Ik Ahn, Seung Min Kim, Gyoo Gun Lim (Hanyang University)

Online customer reviews have become a critical information source in e-commerce, yet the proliferation of reviews has intensified negativity bias, whereby negative reviews exert a disproportionately large influence on purchasing decisions. This study draws on signaling theory to examine whether trust signals, specifically repurchase reviews and one-month usage reviews, are associated with a reduced likelihood of negative review occurrence. Using panel data comprising 732,724 observations across 376 experience goods from a major South Korean e-commerce platform, the study employs a Generalized Linear Mixed-Effects Model (GLMER) to account for product-level heterogeneity. Results indicate that the accumulation of one-month usage reviews is associated with a 1.8 percent reduction in negative review risk, while repurchase reviews are associated with a 2.9 percent reduction. These findings suggest that trust signals function as credible defensive mechanisms against negative review occurrence in e-commerce.

KEY WORDS

Signaling theory, Negativity bias, Repurchase review, One-month usage review, Consumer behavior

KITS·ICEC·KIISS
2026 통합 국제학술대회

B4 Smart Tourism & Immersive Experience

Session Chair : Soyeon Kwon (Dongguk Univ.)



B4

Development of Hybrid Slot-Filling Travel Recommendation System: Focusing on the Runner-Pacer Dual Framework.

Jae Ik Ahn, Su Kyeom Kim, Gyoo Gun Lim (Hanyang University)

Recent travel trends are shifting from supplier-centered package tours toward hyper-personalized travel experiences tailored to individual preferences and contexts. However, existing travel recommendation systems based on collaborative filtering, content-based filtering, and reinforcement learning face limitations in recommendation diversity, especially in small and medium-sized cities where data sparsity and cold-start problems are severe. In addition, sequence-based approaches constrain the discovery of novel destinations and diverse travel routes. To address these issues, this study proposes a personalized hybrid recommendation framework using a slot-filling-based travel route recommendation algorithm. The model was empirically applied to travel data from metropolitan areas (Seoul and Busan) and small-to-medium-sized cities (Jeonju, Andong, and Gongju). The proposed framework integrates sequence-based and content-based models, demonstrating practical potential for AI-driven travel assistants while contributing to more balanced tourism demand across regions.

KEY WORDS

Travel Recommendation, Hybrid Slot-Filling, Sequential Recommendation, Hyper-personalization, Smart Tourism

B4

Does Artificial Intelligence Enhance Tourism Economic Resilience? Evidence for Smart Tourism 2.0

Wanjun Gu, Chulmo Koo (Kyung Hee University).

As the core driver of smart tourism 2.0, artificial intelligence (AI) may reshape how urban tourism systems respond to external shocks, yet its effect on tourism economic resilience remains underexplored. Using panel data from 287 Chinese cities over 2010~2023, this study investigates the effect of AI on tourism economic resilience. Results confirm that AI significantly enhances tourism economic resilience, a finding robust across all identification strategies. This effect operates through smart investment and innovation talent agglomeration, with regional information intensity and tertiary sector share serving as negative moderators. Heterogeneity analyses indicate that effects are most pronounced in cities with higher government digitalization, greater tourism dependency, and resource-city status, and AI-driven tourism economic resilience further reduces employment volatility and strengthens fiscal resilience. This study provides the first empirical evidence for AI-powered smart tourism 2.0 and offers targeted policy recommendations for smart tourism city governance.

KEY WORDS

Smart tourism 2.0, Artificial intelligence, Tourism economic resilience, Smart tourism cities governance

Does Artificial Intelligence Enhance Tourism Economic Resilience? Evidence for Smart Tourism 2.0

Wanjun,Gu.[a] *and Chulmo, Koo.*[b]
[a] *Smart Tourism Education Platform, Kyung Hee University*
E-mail: wjgu97@khu.ac.kr
[b] *Smart Tourism Education Platform, Kyung Hee University*
E-mail: helmetgu@khu.ac.kr

Abstract

As the core driver of smart tourism 2.0, artificial intelligence (AI) may reshape how urban tourism systems respond to external shocks, yet its effect on tourism economic resilience remains underexplored. Using panel data from 287 Chinese cities over 2010–2023, this study investigates the effect of AI on tourism economic resilience. Results confirm that AI significantly enhances tourism economic resilience, a finding robust across all identification strategies. This effect operates through smart investment and innovation talent agglomeration, with regional information intensity and tertiary sector share serving as negative moderators. Heterogeneity analyses indicate that effects are most pronounced in cities with higher government digitalization, greater tourism dependency, and resource-city status, and AI-driven tourism economic resilience further reduces employment volatility and strengthens fiscal resilience. This study provides the first empirical evidence for AI-powered smart tourism 2.0 and offers targeted policy recommendations for smart tourism city governance.

Keywords: Smart tourism 2.0, Artificial intelligence, Tourism economic resilience, Smart tourism cities governance

Introduction

Over the past decade, the rapid advancement of artificial intelligence (AI) has transformed how tourism systems operate, create value, and engage with travelers. This transformation has driven a fundamental transition from smart tourism 1.0 to smart tourism 2.0, reflecting a new era of AI-driven tourism development (Gretzel et al. 2015; Koo et al. 2025). Koo et al. (2025) proposed the "AI-powered Smart Tourism 2.0 model," repositioning AI as the integrating force that reshapes destination management, value co-creation, and the broader tourism business ecosystem. Critically, this model suggests that smart tourism 2.0 not only enhances destination competitiveness and tourist experiences, but also strengthens the governance, adaptive, and responsive capacities of urban tourism systems.

These capacities are especially important because tourism is shaped not only by internal development but increasingly by external shocks, including climate change (Scott and Gössling, 2022), natural disasters (Assaf et al. 2021), and geopolitical conflicts (Papagianni et al. 2023). In China, for example, the COVID-19 pandemic caused domestic tourism revenue to collapse from CNY 5.73 trillion in 2019 to CNY 2.23 trillion in 2020, a decline of 61.1%, while domestic tourist trips fell by over 52% within the same year (Chen et al. 2024). The Russia-Ukraine war similarly exposed geopolitical vulnerabilities, with Ukraine alone recording cumulative tourism revenue losses of USD 19.6 billion within two years of the conflict (ET TravelWorld, 2024). In this context, tourism resilience has

attracted growing attention from governments, international organizations, and the academic community worldwide (Gretzel and Scarpino-Johns, 2018).

A growing body of research has pointed the importance of tourism resilience (Prayag et al. 2023), highlighting its role in sustaining destination competitiveness and mitigating the long-term economic consequences of external shocks. Smart tourism 2.0 explicitly emphasizes resilience as a critical governance requirement, particularly given the social disruptions that accompany the rapid rise of the AI-powered economy (Gretzel and Scarpino-Johns, 2018; Koo et al. 2025). However, the conceptualization of tourism resilience remains contested, with different studies examining tourism economic resilience (Cheng and Zhang, 2020), tourism ecological resilience (Ruiz-Ballesteros, 2011), and tourism organizational resilience (Prayag et al. 2023) , among others. This study focuses on tourism economic resilience, as it most directly captures the governance-relevant performance of urban tourism systems under external shocks. Drawing on Lu (2022), this study defines tourism economic resilience as the capacity of a tourism economic system to resist external shocks, recover and adjust its core functions, and reorganize and upgrade its development trajectory in the face of disruption. Tourism economic resilience therefore provides a particularly suitable perspective for assessing the practical implications of smart tourism 2.0.

As the core driver of smart tourism 2.0, AI provides a concrete point of entry for assessing the framework's practical implications. Existing research has widely shown that AI can promote resilience across multiple domains, including firm resilience (Han et al. 2025), energy resilience (Nepal et al. 2024), and supply chain resilience (Lu et al. 2025). In tourism, existing studies have shown that digitalization and intelligent technologies can contribute to tourism economic resilience (Tang, 2023; Lu et al. 2024). However, the digital economy broadly captures the general digital environment and infrastructure, whereas AI represents a more specific and targeted technological driver. Moreover, the core transformation from smart tourism 1.0 to smart tourism 2.0 is precisely the technological shift from digitalization to AI. This distinction matters because the resilience implications of AI cannot be adequately reflected by studies on the digital economy or digital technologies in general. Yet the relationship between AI and tourism economic resilience has remained underexplored. By addressing this gap, this study provides clearer empirical evidence for smart tourism 2.0, using tourism economic resilience as the analytical perspective and AI as the empirical entry point. Thus, we propose the following research questions:

RQ1: Does AI enhance tourism economic resilience?

RQ2: If so, what are the underlying mechanisms and economic consequences of this effect?

To address these questions, this study uses panel data from 287 Chinese cities over the period 2010–2023 to empirically examine the effect of AI on tourism economic resilience. Theoretically, this study provides the first empirical evidence that AI enhances urban tourism economic resilience, advancing smart tourism 2.0 from conceptual vision toward empirical grounding. Practically, the findings offer policymakers evidence-based guidance for leveraging AI as a core governance instrument in smart tourism city development.

Research Method

2.1. Variables

2.1.1. Dependent variable

Existing studies have mainly measured tourism economic resilience through formula-based approaches (Tang, 2024; Lu, 2022). However, such methods typically rely on a single economic indicator, failing to capture the multidimensional and dynamic nature of resilience. Drawing on Wang (2025), this study constructs a comprehensive evaluation indicator system comprising three dimensions of resistance, recovery, and renewal with 14 indicators in total. These three dimensions are grounded in the evolutionary resilience model (Martin, 2012), where resistance marks the initial phase of withstanding external shocks, recovery represents the intermediate phase of functional restoration, and renewal constitutes the advanced phase of structural reorganization and upgrading. Together, they form a complete resilience chain. This approach more fully captures the conceptual

richness of tourism economic resilience than single-formula measures. We employ the entropy weight method to calculate the composite index, which avoids subjective judgment and transforms heterogeneous indicators into dimensionless information entropy values.

2.1.2. Independent variable

To measure city-level AI development, this study follows Wang (2022) and uses the number of AI firms in each city as the proxy variable. The number of AI firms reflects the local agglomeration of AI-related technological activity and innovation capacity, thereby is a reliable indicator of city-level AI development. Such AI development can further spill over into the tourism system through digital infrastructure and intelligent services. We identify AI firms through a keyword-based feature recognition approach, using the registered business scope information publicly disclosed in the Tianyancha enterprise database as the textual basis for analysis. The business scope of AI firms is defined by keywords related to artificial intelligence, including chips, image recognition, computer vision, speech recognition, and sensors. To address the issue of cities with no AI firms, the independent variable is further specified as ln (AI+1) to better reflect city-level AI development.

2.1.3. Other variables

Following existing studies (Bhati et al. 2016; Prodi et al. 2023; P.L. et al. 2021), this study selects six control variables that may affect urban tourism economic resilience, including gross domestic product, industrial structure level, government regulation level, financial development level, transportation infrastructure level, and import-export level. Mediating variables include city-level smart investment and innovative talent agglomeration. Moderating variables include regional tertiary sector share and regional information intensity, reflecting the external conditioning factors between AI and tourism economic resilience. Economic consequence variables include employment stability and fiscal resilience.

2.2. Model setting

To examine the effect of AI on urban tourism economic resilience, this study employs a two-way fixed effects model as follows:

$$TER_{it} = \beta_0 + \beta_1 ln(AI+1)_{it} + \beta_2 Controls_{it} + \mu_i + \nu_t + \varepsilon_{it} \quad (1)$$

Furthermore, to address RQ2 regarding the mechanisms and economic consequences of AI's effect on tourism economic resilience, this study constructs the following extended models.

$$Med_{it} = \alpha_0 + \alpha_1 ln(AI+1)_{it} + \alpha_2 Controls_{it} + \mu_i + \nu_t + \varepsilon_{it} \quad (2)$$

$$TER_{it} = \gamma_0 + \gamma_1 ln(AI+1)_{it} \times Mod_{it} + \gamma_2 ln(AI+1)_{it} + \gamma_3 Mod_{it} + \gamma_4 Controls_{it} + \mu_i + \nu_t + \varepsilon_{it} \quad (3)$$

$$Conseq_{it} = \lambda_0 + \lambda_1 ln(AI+1)_{it} \times TER_{it} + \lambda_2 ln(AI+1)_{it} + \lambda_3 TER_{it} + \lambda_4 Controls_{it} + \mu_i + \nu_t + \varepsilon_{it} \quad (4)$$

where i and t represent the city and the year, respectively; TER_{it} is the dependent variable; $\ln (AI+1)_{it}$ denotes the independent variable; Med_{it} is the Mediator variables; Mod_{it} is the moderator variables; $Conseq_{it}$ denotes the economic consequences; $Controls_{it}$ refers to the series of control variables; ε_{it} represents the random error term; β_1 is the coefficient of the independent variable; μ_i and ν_t represent the fixed effects for city and year respectively. Additionally, standard errors are clustered at the city level to ensure robustness.

Empirical Results

3.1 Descriptive statistics

Table 1 presents the descriptive statistics of the variables in the research model. Tourism economic resilience shows a mean of 0.076 and a standard deviation of 0.086, with values ranging from a minimum of 0.010 to a maximum of 0.971. The primary independent variable AI has a mean of 1.093

and a standard deviation of 1.360, spanning from 0 to 7.026. These statistical measures suggest notable variations in the level of AI development across Chinese cities.

Table 1. Descriptive Statistics.

Variables	Obs	Mean	Std. Dev.	Min	Max
TER	3214	.076	.086	.010	.971
Ln(AI+1)	3214	1.093	1.360	.000	7.026
LnGDP	3214	23.646	.993	21.014	26.881
Inds	3214	1.060	.604	.109	5.691
Gov	3214	.194	.100	.044	1.182
Fin	3214	25.904	1.279	22.795	30.016
Croad	3214	2.886	.406	.863	4.112
Ei	3214	.404	.054	.027	.648
Si	3214	1.048	1.236	.000	3.871
Talent	3214	.052	.085	.002	.499
Thirdind	3214	1.041	.508	.373	2.862
Internet	3214	13.767	1.149	10.061	17.332
Employ	3214	.228	.372	.025	3.449
Fiscal	3214	.224	.319	.000	1.074

3.2. Benchmark results

To examine the effect of AI on tourism economic resilience, we conduct baseline regression analyses. The results are presented in Table 2, Column (1) shows the results without control variables, while Column (4) is the full model settings with all controls included. Across all estimated models, the coefficient of AI remains statistically significant at the 1% level. Specifically, a one-standard-deviation increase in AI is associated with a 1.02% (0.0139 ÷ 1.360) increase in tourism economic resilience, confirming that AI significantly enhances tourism economic resilience and answering RQ1.

Table 2. Benchmark Regression.

Variable	Dependent variable: Tourism economic resilience			
	(1)	(2)	(3)	(4)
Ln(AI+1)	0.0123***	0.0127***	0.0127***	0.0139***
	(0.0040)	(0.0042)	(0.0043)	(0.0045)
LnGDP		-0.0099	-0.0284**	-0.0302**
		(0.0125)	(0.0130)	(0.0131)
Inds		-0.0040	-0.0051	-0.0065
		(0.0112)	(0.0118)	(0.0115)
Gov			-0.0603	-0.0409
			(0.0488)	(0.0500)
Fin			0.0261**	0.0242**
			(0.0106)	(0.0104)
Croad				0.0299***

				(0.0115)
Ei				0.1021**
				(0.0395)
Constant	0.0621***	0.3008	0.0744	0.0357
	(0.0043)	(0.3002)	(0.3119)	(0.3034)
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3,214	3,214	3,214	3,214
R-squared	0.727	0.727	0.729	0.734

Note: Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. To account for potential heteroskedasticity and serial correlation in the sample data, standard errors are clustered at the city level.

3.3. Robustness check

To verify the robustness of the findings, seven identification strategies are employed, including Bartik IV, propensity score matching (1:1, 1:2, 1:3), placebo tests, omitted variable sensitivity analysis, policy controls, sample exclusions, and alternative clustering dimensions. Across all these checks, the positive effect of AI on tourism economic resilience remains significant, confirming the robustness of the main findings.

3.4 Further analysis

To further address RQ2, this section examines how AI enhances tourism economic resilience through specific mechanisms, under what conditions these effects are amplified or dampened, and what broader economic consequences follow.

Mediation analyses reveal that AI enhances tourism economic resilience via two pathways: promoting smart investment and agglomerating innovation talent. Moderation analyses further indicate that both regional information intensity and tertiary sector share significantly attenuate this positive effect. In cities with well-established information infrastructure or higher service sector maturity, the marginal returns of AI diminish as the structural space for AI-driven resilience gains narrows. Consequence models additionally demonstrate that AI-driven improvements in tourism economic resilience generate broader macroeconomic spillovers, effectively stabilizing urban labor markets and strengthening local fiscal resilience (see in table 3).

Table 3. Further Analysis Results.

	Mediation		Moderation		Consequence	
Variable	SI	Talent	TER	TER	Employ	Fiscal
TER					-0.2719**	0.2664
					(0.1246)	(0.1794)
AI	0.8946***	0.0048***	0.1431***	0.0330***	0.0121***	-0.0766***
	(0.0229)	(0.0012)	(0.0237)	(0.0071)	(0.0046)	(0.0093)
Internet			0.0244***			
			(0.0086)			
AI*Internet			-0.0088***			
			(0.0016)			
Thirdind				0.0530*		

				(0.0297)		
AI*Thirdind				-0.0136***		
				(0.0041)		
TER*AI					0.0661*	0.2622***
					(0.0342)	(0.0815)
Constant	2.5990**	-0.4046***	-0.0750	-0.3148	-1.6831***	6.9374***
	(1.1736)	(0.1286)	(0.2535)	(0.2403)	(0.4950)	(0.6835)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,214	3,214	3,214	3,214	3,214	3,214
R-squared	0.982	0.980	0.746	0.744	0.982	0.904

Note: Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. To account for potential heteroskedasticity and serial correlation in the sample data, standard errors are clustered at the city level.

Heterogeneity analyses reveal that AI's positive effect on tourism economic resilience is stronger in cities with higher government digitalization, greater tourism dependence, and resource-based industrial structures (see in table 4).

Table 4. Heterogeneity Test.

Variable	Government Digitalisation		Tourism Dependence		Resource Cities	
	Low	High	High	Low	No	Yes
AI	0.0266**	0.0051***	0.0154***	0.0016	0.0077***	0.0158***
	(0.0105)	(0.0019)	(0.0050)	(0.0055)	(0.0012)	(0.0059)
Constant	1.1803	-0.3960*	-0.3037	0.5478	-0.5317***	0.1922
	(0.8139)	(0.2189)	(0.3575)	(0.4723)	(0.1409)	(0.5696)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	917	2,297	2,546	664	1,195	1,988
R-squared	0.750	0.720	0.757	0.679	0.781	0.746

Note: Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. To account for potential heteroskedasticity and serial correlation in the sample data, standard errors are clustered at the city level.

Conclusion and Discussion

4.1. Conclusion

Using panel data from 287 Chinese cities over 2010–2023, this study confirms that AI significantly enhances urban tourism economic resilience. This finding remains robust across seven identification strategies, including Bartik instrumental variables. Mechanism analyses reveal that this effect operates through Smart investment and innovation talent agglomeration, with higher tertiary sector shares and

regional information intensity moderating the strength of this effect. Further analysis demonstrates that AI-driven tourism economic resilience reduces employment volatility and strengthens urban fiscal resilience. Heterogeneity analyses show that these effects are particularly evident in cities with higher government digitalization, greater tourism dependency, and resource city status.

4.2. Theoretical implications

This study makes a distinctive theoretical implications to the smart tourism 2.0 framework by moving it from a conceptual proposition toward an empirically verifiable governance framework. Specifically, we demonstrate that AI significantly enhances tourism economic resilience, providing the first robust empirical evidence for AI-powered smart tourism 2.0 We also extend the AI-powered smart tourism 2.0 model by establishing the connection between the AI layer and the physical layer. Beyond this primary finding, the economic consequence analyses address a specific theoretical question within smart tourism 2.0 concerning the relationship between the experience layer and the physical layer. The finding that AI-driven tourism economic resilience stabilizes urban fiscal revenues and reduces employment volatility extends the literature on AI and smart city governance (Wolniak and Stecuła, 2024; Kim et al. 2025; Gretzel et al. 2016).

This study also contributes to the tourism economic resilience literature. Existing studies have mainly examined the influence of factors such as the digital economy and digital transformation on tourism economic resilience (Tang, 2024; Lu et al. 2025), leaving AI as its most representative component largely unexplored. To the best of our knowledge, this study represents the first empirical investigation into the relationship between AI and tourism economic resilience, enriching the tourism economic resilience literature. More importantly, tourism economic resilience is not merely a passive outcome of external shocks, but a governance-oriented capacity that can be actively shaped through AI-driven transformation. In doing so, this study moves tourism resilience research beyond the traditional "shock-recovery" logic toward a new paradigm of "technology empowerment and resilience formation" (Wei et al. 2025). This suggests that tourism economic resilience can be understood not only as an outcome of shocks, but also as an expression of Smart Tourism 2.0 governance capacity.

4.3. Policy recommendations

Based on the findings of this study, the following policy recommendations are proposed for smart tourism city governance. First, given that AI significantly enhances urban tourism economic resilience suggests that governments should not treat AI merely as a service tool for tourist experience enhancement. Instead, AI should be a core governance instrument for strengthening the resilience of urban tourism systems. Second, the mediating analyses reveal that this effect operates through smart investment and innovation talent agglomeration. Policy design should therefore focus less on AI deployment in isolation. Smart infrastructure construction and innovation talent cultivation should be advanced simultaneously, thereby more effectively translating AI into tourism resilience gains. Finally, the heterogeneity analyses show that AI's resilience-enhancing effects are strongest in cities with higher government digitalization, greater tourism dependency, and resource-city status. This suggests that Smart Tourism 2.0 should be implemented through differentiated strategies rather than a one-size-fits-all approach. It is recommended that pilot zones be established as a priority in cities exhibiting these characteristics, where AI-driven governance models can be tested, refined, and subsequently scaled to similar cities nationwide.

References

Assaf, A. G., Kock, F., and Tsionas, M. 2021. “Tourism during and after COVID-19: An Expert-Informed Agenda for Future Research,” *Journal of Travel Research* (61:2), 004728752110172.

Bhati, A., Upadhayaya, A., and Sharma, A. 2016. “National Disaster Management in the ASEAN-5: An Analysis of Tourism Resilience,” *Tourism Review* (71:2), pp. 148–164.

Chen, Y., Wu, F., Zhang, D., and Ji, Q. 2024. “Tourism in Pandemic: The Role of Digital Travel Vouchers in China,” *Humanities and Social Sciences Communications* (11:1), Palgrave Macmillan, pp. 1–15.

Cheng, L., and Zhang, J. 2020. "Is Tourism Development a Catalyst of Economic Recovery Following Natural Disaster? An Analysis of Economic Resilience and Spatial Variability," *Current Issues in Tourism* (23:20), pp. 2602–2623.

ET TravelWorld. 2024. "UNESCO Says $9 Billion Needed to Revive Ukraine Tourism," *ETTravelWorld.com*, Reuters, February 14.

Gretzel, U., and Scarpino-Johns, M. 2018. "Destination Resilience and Smart Tourism Destinations," *Tourism Review International* (22:3), pp. 263–276.

Gretzel, U., Sigala, M., Xiang, Z., and Koo, C. 2015. "Smart Tourism: Foundations and Developments," *Electronic Markets* (25:3), pp. 179–188.

Gretzel, U., Zhong, L., and Koo, C. 2016. "Application of Smart Tourism to Cities," *International Journal of Tourism Cities* (2:2).

Han, M., Shen, H., Wu, J., and Zhang, X. (Michael). 2025. "Artificial Intelligence and Firm Resilience: Empirical Evidence from Natural Disaster Shocks," *Information Systems Research* (36:4), pp. 2116–2133.

Kim, J. Y., Gretzel, U., Berezina, K., Alzua-Sorzabal, A., and Koo, C. 2025. "TechTalk Interview about Smart Tourism 2.0: Setting the Research Agenda," *Journal of Smart Tourism* (5:3), pp. 73–81.

Koo, C., Xiang, Z., Gretzel, U., and Sigala, M. 2021. "Artificial Intelligence (AI) and Robotics in Travel, Hospitality and Leisure," *Electronic Markets* (31:3), p. 473.

Lu, X., Chiang, C.-Y., Wu, F., and Xu, X. 2025. "Whose AI Matters? Examining the Bilateral Effects of AI Capability Orientation on Supply Chain Resilience," *International Journal of Operations & Production Management* (45:8), Emerald, pp. 1549–1577.

Lu, X., Zhu, J., Zhou, M., and Zhang, Y. 2024. "Is Digital Economy the Driving Force for Improving the Tourism Economic Resilience? Evidence from China," *Environment, Development and Sustainability* (27:12), pp. 30249–30286.

Martin, R. 2012. "Regional Economic Resilience, Hysteresis and Recessionary Shocks," *Journal of Economic Geography* (12:1), pp. 1–32.

Nepal, R., Zhao, X., Dong, K., Wang, J., and Sharif, A. 2024. "Can Artificial Intelligence Technology Innovation Boost Energy Resilience? The Role of Green Finance," *Energy Economics* (142), North-Holland, p. 108159.

P.L., B., Tzu-Ling Chen, C., and Wickens, E. 2021. "Tourism Industry Resilience Issues in Urban Areas during COVID-19," *International Journal of Tourism Cities* (7:3), pp. 861–879.

Papagianni, E., Evgenidis, A., Τσαγκανός, A., and Megalooikonomou, V. 2023. "Tourism Demand in the Face of Geopolitical Risk: Insights from a Cross-Country Analysis," *Journal of Travel Research* (63:8), SAGE Publishing, pp. 2094–2119.

Prayag, G., Jiang, Y., Chowdhury, M., Hossain, M. I., and Akter, N. 2023. "Building Dynamic Capabilities and Organizational Resilience in Tourism Firms during COVID-19: A Staged Approach," *Journal of Travel Research* (63:3), 004728752311649-004728752311649.

Prayag, G., Muskat, B., and Dassanayake, C. 2023. "Leading for Resilience: Fostering Employee and Organizational Resilience in Tourism Firms," *Journal of Travel Research* (63:3), pp. 659–680.

Prodi, E., Fasone, V., and Di, M. R. 2023. "Does Industry Resilience Matter for Postshock Industrial Policy? A Focus on Tourism-Related Industries," *Tourism Economics* (30:2), SAGE Publishing, pp. 389–416.

Ruiz-Ballesteros, E. 2011. "Social-Ecological Resilience and Community-Based Tourism," *Tourism Management* (32:3), pp. 655–666.

Scott, D., and Gössling, S. 2022. "A Review of Research into Tourism and Climate Change - Launching the Annals of Tourism Research Curated Collection on Tourism and Climate Change," *Annals of Tourism Research* (95), p. 103409.

Tang, R. 2023. "Can Digital Economy Improve Tourism Economic Resilience? Evidence from China," *Tourism Economics* (30:6), SAGE Publishing, pp. 1359–1381.

Wei, N., Xie, W.-J., and Zhou, W.-X. 2025. "Resilience of International Oil Trade Networks under Extreme Event Shock-Recovery Simulations," *Energy* (314), Elsevier BV, p. 134174.

Wolniak, R., and Steculа, K. 2024. "Artificial Intelligence in Smart Cities—Applications, Barriers, and Future Directions: A Review," *Smart Cities* (7:3), pp. 1346–1389.

B4

Framing Scarcity: The Joint Effects of Availability Quantity and Presentation Format on Consumer Decisions

Danhye Kyoung, Chulmo Koo (KyungHee University)

With the rapid growth of OTAs, effective online strategies for selling perishable goods like hotel rooms and airline tickets have become increasingly important. This study examines how decision frames in availability information–specifically, presentation format (rate vs. percentage) and availability display (remain vs. sold)–influence consumer decision–making through a 2×2 scenario experiment.

KEY WORDS

OTAs, framing effect, Scarcity Effect, Processing Fluency

Indicate Submission Type: Framing Scarcity: The Joint Effects of Availability Quantity and Presentation Format on Consumer Decisions

Kyoung, D. H.[a] *and Koo, C. M.*[b]
[a] *Smart Tourism Education Platform (STEP), Kyung Hee University*
E-mail: danhye9298@khu.ac.kr
[b] *Smart Tourism Education Platform (STEP), Kyung Hee University*
E-mail: helmetgu@khu.ac.kr

Abstract

This study investigates how different forms of availability framing influence consumer decision-making in online travel agencies (OTAs), where selling perishable products such as hotel rooms requires effective information design strategies. While prior research has extensively examined scarcity effects, limited attention has been paid to how the format of scarcity information shapes consumer responses. Addressing this gap, the present study examines the joint effects of availability framing (remain vs. sold) and numerical presentation format (Rate vs. percentage) on perceived risk and purchase attitude. Using a two-study experimental design, Study 1 validates the scarcity effect, while Study 2 employs a 2×2 between-subjects experiment to test the interaction between framing and numerical format. The results reveal no significant main effects; however, a significant interaction effect emerges. Specifically, in the "sold" condition, percentage formats increase perceived risk and reduce purchase attitude compared to absolute numerical formats. In contrast, in the "remain" condition, percentage formats enhance purchase evaluation relative to numerical counts. Furthermore, moderated mediation analysis (PROCESS Model 7) demonstrates that perceived risk mediates this interaction effect, with significant indirect effects observed only under the "sold" framing. These findings extend framing and information processing theories by demonstrating that the effectiveness of scarcity cues depends not only on the level of scarcity but also on how numerical information is presented. From a practical perspective, the results suggest that OTA platforms can optimize consumer responses by strategically aligning availability framing with appropriate numerical formats, thereby reducing perceived risk and improving purchase outcomes..

Keywords: OTAs, framing effect, Scarcity Effect, Processing Fluency

Introduction

As of 2024, 82% of global bookings are made online, with OTAs accounting for 41% of the market (Yang, 2023). Consumers increasingly trust online platforms due to price transparency and ease of comparison, making OTA sales strategies critical for success. Previous studies have examined psychological marketing tactics and consumer traits like impulsivity, with particular attention to how scarcity effects drive purchase behavior (Kim et al., 2021). Building on this foundation, the present study (1) reevaluates the effectiveness of scarcity and (2) investigates how different framing formats (remain vs. sold; rate vs. percentage) influence perceived risk and purchase behavior, based on decision

framing theory. Study 1 reviews the theoretical background and validates the scarcity effect, while Study 2 empirically tests the impact of availability and numerical framing in a hotel booking context. The findings provide practical implications for OTA marketers and advance theoretical insights into consumer decision-making in digital tourism environments.

2. Literature Review and Hypothesis Development

On each page, your material (not including the header and footer) should fit within a rectangle of 210 × 297 mm (8.27 × 11.69 in.), centered on an A4 page, beginning 1.9 cm (.75 in.) from the top of the page. Please adhere to the A4 size only (hopefully Word or other word processors can help you with it). If you cannot do so, please contact the review coordinator for assistance. All final publications will be formatted and displayed in A4 size. Right margins should be justified, not ragged. All margins must measure 1" (2.5 cm) around. Beware, especially when using this template on a Macintosh, Word may change these dimensions in unexpected ways.

2.1. Scarcity Effect

Scarcity effect suggests that when a product or resource is perceived as limited, consumers assign it greater value and feel a stronger urgency to act (Cialdini, 2001). In online booking environments, scarcity cues such as "Only a few left" or "Selling fast" are commonly used to create purchase pressure and accelerate decision-making. Thus, the following hypothesis is proposed:

H1. When the available quantity is low (vs. high), consumers show higher purchase attitudes.

2.2. Framing Effect and Information Processing Theory

Framing Effect Theory suggests that identical information can lead to different consumer attitudes and behaviors depending on how it is presented (Tversky & Kahneman, 1981). In online booking contexts, availability can be framed as either remaining quantity (e.g., "10 rooms left") or sold quantity (e.g., "90 rooms sold"), each triggering different psychological perceptions—gain vs. loss. While loss-framed messages often heighten urgency and influence decisions, they can also backfire, especially when combined with unclear numerical formats.

Beyond framing, the format of numerical presentation—rate (e.g., "10 rooms") vs. percentage (e.g., "10%")—affects how consumers process information. According to Information Processing Theory (Alba et al., 1987; Eagly & Chaiken, 1993), less intuitive formats like percentages increase cognitive load, particularly when reference points are ambiguous (Peters et al., 2006). This can reduce processing fluency, elevate perceived risk, and lower purchase intention.

Based on this theoretical background, the present study proposes that the interaction between availability presentation (Remain vs. Sold) and presentation format (Rate vs. %) will have a significant effect. Specifically, the following hypothesis is proposed:

H2. The effect on consumers' purchase attitude varies depending on the combination of availability presentation (Remain vs. Sold) and presentation format (Rate vs. Percentage).

H2a. When availability is presented as Sold (vs. Remain) in percentage format (vs. rate), it negatively affects consumers' purchase attitude.

H2b. When availability is presented as Remain (vs. Sold) in rate format (vs. percentage), it negatively affects consumers' purchase attitude.

H3. Perceived risk mediates the relationship between availability framing and purchase intention.

Figure 1 represents the research model based on our hypothesis.

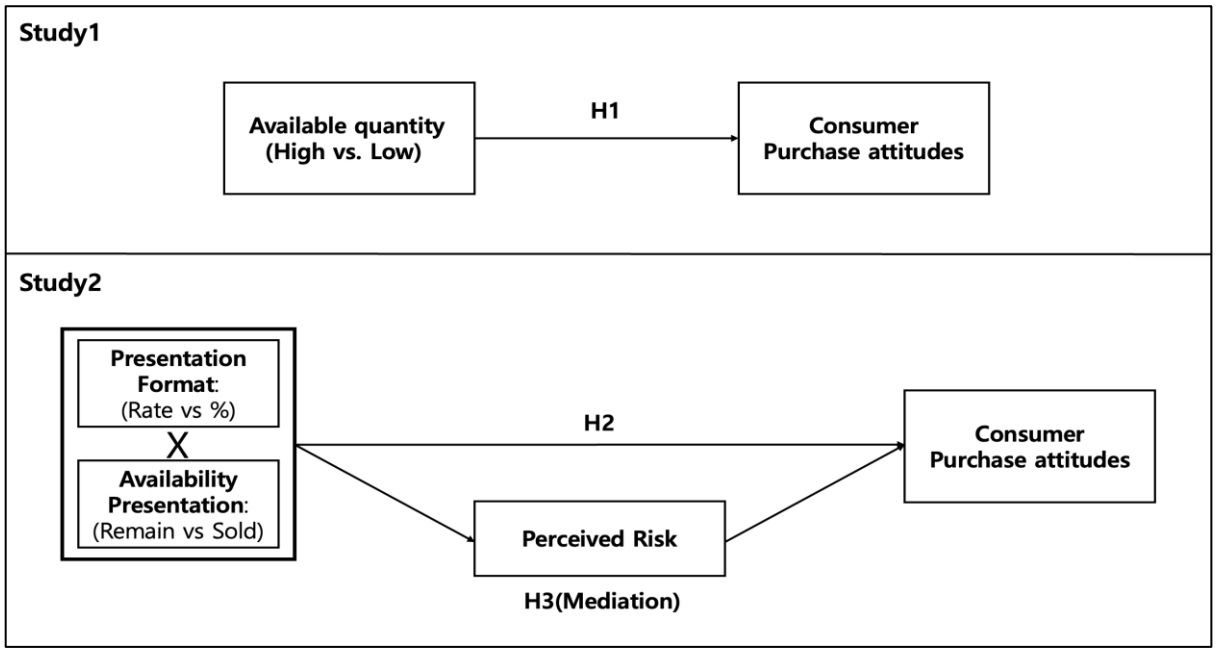


Figure 1. Research Model

3. Methodology

This study conducted two experiments to test the proposed hypotheses. Study 1 aimed to examine the scarcity effect, while Study 2 employed a 2 × 2 experimental design to investigate the effects of availability framing (sold vs. remain) and numerical presentation format (Rate vs. percentage). A total of 306 U.S.-based online panel participants were recruited (Study 1: n = 102; Study 2: n = 205), and participants were randomly assigned to one of four experimental conditions. After being exposed to a hypothetical scenario and experimental stimuli (Figure 2), participants evaluated their perceived risk and purchase attitude. All measures were assessed using a 7-point Likert scale.

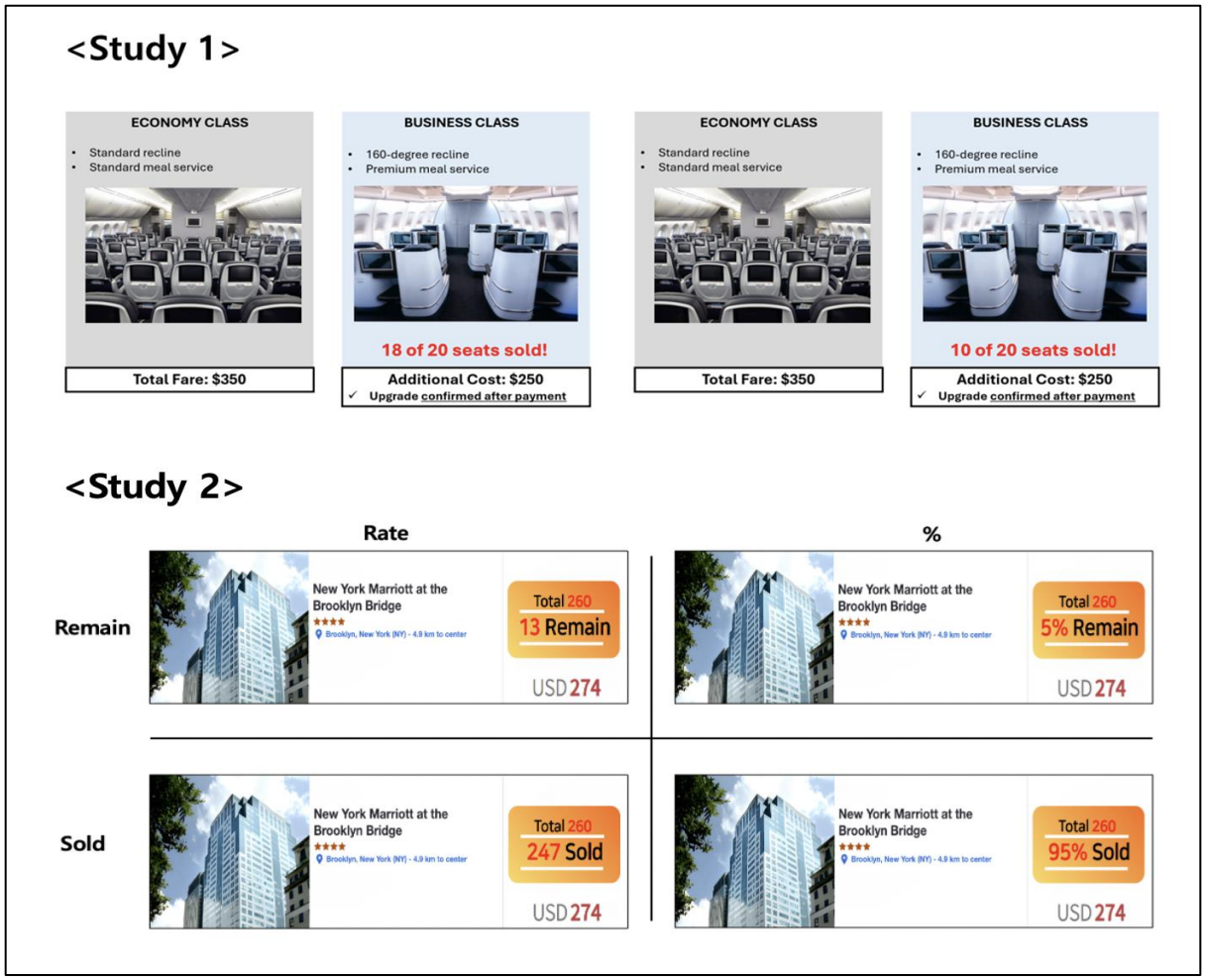


Figure 2. Stimuli for Study 1 and Study 2

4. Results

4.1. Study1

A manipulation check confirmed that participants perceived the scarcity level as intended. Those in the high scarcity condition ("18/20 sold") reported significantly higher perceived scarcity ($M = 6.02$) than those in the low scarcity condition ("10/20 sold"; $M = 4.45$; $t(100) = 6.01$, $p < .001$). A one-way ANOVA was conducted to examine the effect of scarcity level on consumers' purchase attitude. Participants in the high scarcity condition reported a significantly higher purchase attitude ($M = 4.29$, $SD = 1.81$) compared to those in the low scarcity condition ($M = 3.51$, $SD = 1.63$). The analysis revealed a significant main effect of scarcity level on purchase attitude ($F(1, 99) = 4.76$, $p = .032$, $\eta^2 = .046$). These results suggest that greater perceived scarcity—as indicated by the "18/20 sold" message—enhances consumers' attitude toward purchase.

4.2. Study2

A manipulation check confirmed that participants correctly identified their assigned conditions ($F(1,203) = 172.63$, $p < .001$, $\eta^2 = .460$) and rated the scenario as realistic ($t(204) = 16.53$, $p < .001$, $M = 5.40$). A 2×2 ANOVA examined the effects of availability framing (Sold vs. Remain) and numeric format (% vs. Rate) on perceived risk and purchase attitude. No significant main effects were found for either framing ($F = .940$, $p = .334$) or format ($F = .651$, $p = .421$). However, the interaction effect was significant for perceived risk ($F = 7.712$, $p = .006$) and purchase attitude ($F = 7.529$, $p = .007$). In the "Sold" condition, percentage format increased perceived risk ($M = 3.66$) and reduced purchase intention ($M = 3.85$) compared to rate format. In the "Remain" condition, the opposite pattern was found: rate

format increased perceived risk (M = 3.69), while percentage format slightly increased purchase intention (M = 4.79). A moderated mediation analysis (PROCESS Model 7) showed that perceived risk mediated the effect of numeric format on purchase attitude, moderated by framing, with significant indirect effects only in the "Sold" condition.

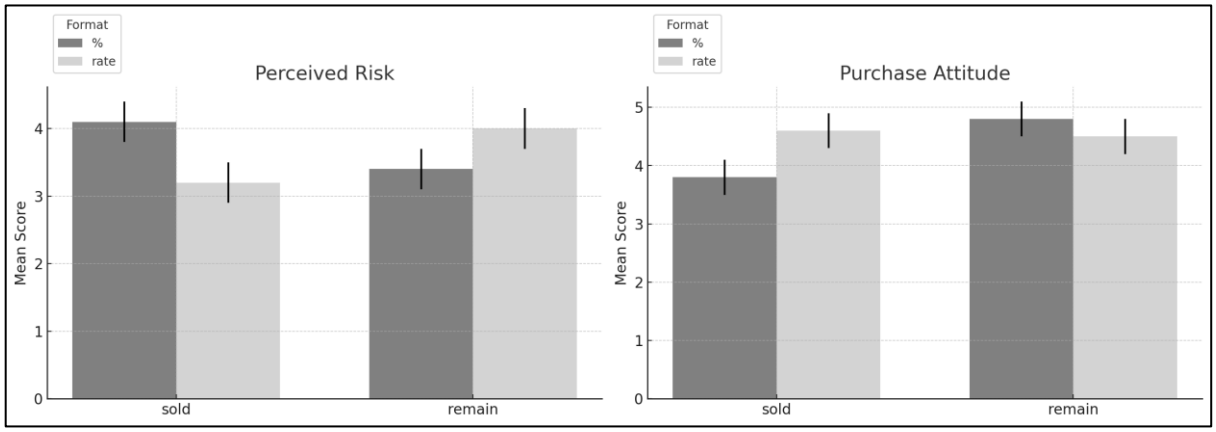


Figure 3. Results of Study2

5. Discussion and Implications

5.1. Discussion

This study showed that in the “Sold” condition, quantity-based (rate) information heightened urgency, strongly influencing consumer behavior. In contrast, the “Remain” condition had a greater impact with percentage (%) information, highlighting the decision-framing effect. The findings emphasize that while scarcity is important, the way it is framed plays a crucial role in shaping consumer responses, extending the discussion from scarcity effects to the influence of different scarcity cues.

5.2. Academic and practical implications

This study highlights the impact of numerical presentation formats in availability framing, unlike prior research focused on price or visuals. It suggests that OTAs can enhance sales and consumer decision-making by displaying both total and remaining quantities of rooms or tickets through effective information design.

5.3. Limitations of this study and suggestions for future research

Future research should broaden the scope to include a wider range of OTA products, such as tours and rental cars. Additionally, incorporating factors such as booking lead time, level of product information exposure, and perceived quality may allow for the development of more sophisticated and targeted marketing strategies.

References

Alba, J. W., Hutchinson, J. W., & Lynch, J. G. (1987). Memory and decision making. In J. Sheth & A. Mitra (Eds.), *Advances in Consumer Research* (Vol. 14, pp. 1–4). Association for Consumer Research.

Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Allyn and Bacon.

Cunningham, S. M. (1967). The major dimensions of perceived risk. In D. F. Cox (Ed.), *Risk taking and information handling in consumer behavior* (pp. 82–108). Harvard University Press.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Hmurovic, J., Lamberton, C., & Goldsmith, K. (2022). Examining the efficacy of time scarcity marketing promotions in online retail. *Journal of Marketing Research, 60*(2), 299–328. (Original work published 2023)

Kim, J. (2021). The effects of price framing strategies on hotel reservation intention in online travel agencies. *Journal of Hospitality and Tourism Research, 45*(3), 410–429.

Kim, J., Jhang, J., Kim, S. S., & Chen, S.-C. (2021). Effects of concealing vs. displaying prices on consumer perceptions of hospitality products. *International Journal of Hospitality Management, 92*, 102708.

Kim, J., Kim, P. B., & Kim, J. E. (2018). Different or similar choices: The effect of decision framing on variety seeking in travel bundle packages. *Journal of Travel Research, 57*(1), 99–115.

Lynn, M. (1991). Scarcity effects on value: A quantitative review of the commodity theory literature. *Psychology & Marketing, 8*(1), 43–57.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Yang, B., Li, S., Chen, Z., & Mattila, A. S. (2023). Consumer responses to time-based sales messages. *Annals of Tourism Research, 99*, 103539.

Zhu, M., & Ratner, R. K. (2015). Scarcity polarizes preferences: The impact on choice among multiple items in a product class. *Journal of Marketing Research, 52*(1), 13–26.

B4

Exploring the Actor-Network Formation of Edu-Metaverse: A Main Path Analysis

Van-Tan Nguyen, Chen-Hao Huang, Ching-Ying Huang
(Department of Business Administration, National Cheng Kung University, Tainan, Taiwan)

Edu-Metaverse research has grown rapidly but remains theoretically fragmented, with studies often examining technologies, pedagogy, and adoption in isolation. This study addresses this gap by integrating Actor-Network Theory (ANT) with Key-Route Main Path Analysis (MPA) to trace the field's dominant development trajectory. Using 3,049 Web of Science publications (2006~2025), the analysis identifies the most influential knowledge path and interprets it through four ANT phases: Problematization, Interessement, Enrolment, and Mobilization. The findings reveal progression from early experimentation with virtual environments, to pedagogical alignment of AR/VR, to stabilized immersive learning systems, and finally to ecosystem-level configurations involving platforms, AI, and governance. The study reconceptualizes Edu-Metaverse as an evolving actor-network and provides a process-oriented explanation of how socio-technical educational systems emerge and stabilize.

KEY WORDS

Edu-Metaverse, Actor-Network Theory, Main Path Analysis, Key-Route MPA, Socio-Technical Systems

Exploring the Actor-Network Formation of Edu-Metaverse: A Main Path Analysis

Nguyen, V. T.[a], Huang, C. H.[b] and Huang, C. Y.[c]
[a] Department of Business Administration, National Cheng Kung University, Tainan, Taiwan
E-mail: ra8137028@gs.ncku.edu.tw
[b] Department of Business Administration, National Cheng Kung University, Tainan, Taiwan
E-mail: chhuang@gs.ncku.edu.tw
[c] Department of Business Administration, National Cheng Kung University, Tainan, Taiwan
E-mail: isastat@mail.ncku.edu.tw

Abstract

Edu-Metaverse research has grown rapidly but remains theoretically fragmented, with studies often examining technologies, pedagogy, and adoption in isolation. This study addresses this gap by integrating Actor-Network Theory (ANT) with Key-Route Main Path Analysis (MPA) to trace the field's dominant development trajectory. Using 3,049 Web of Science publications (2006–2025), the analysis identifies the most influential knowledge path and interprets it through four ANT phases: Problematization, Interessement, Enrolment, and Mobilization. The findings reveal progression from early experimentation with virtual environments, to pedagogical alignment of AR/VR, to stabilized immersive learning systems, and finally to ecosystem-level configurations involving platforms, AI, and governance. The study reconceptualizes Edu-Metaverse as an evolving actor-network and provides a process-oriented explanation of how socio-technical educational systems emerge and stabilize.

Keywords: Edu-Metaverse, Actor-Network Theory, Main Path Analysis, Key-Route MPA, Socio-Technical Systems

Introduction

Edu-Metaverse research has grown rapidly in recent years, reflecting increasing scholarly interest in immersive learning, virtual worlds, digital platforms, and artificial intelligence in education (Dwivedi et al., 2022; Won et al., 2023; Sripan & Jeerapattanatorn, 2025). However, the field remains theoretically fragmented. Existing studies often focus on specific technologies, such as virtual reality, augmented reality, or metaverse platforms, or examine particular outcomes such as engagement, adoption, and learning performance, without clearly explaining how these elements have evolved into a broader educational system (Radianti et al., 2020; Hwang & Chien, 2022; Wu & Hao, 2023).

This limitation is important because Edu-Metaverse is not simply a collection of digital tools. Rather, it involves the ongoing alignment of multiple human and non-human actors, including learners, educators, platforms, immersive devices, intelligent systems, and institutional arrangements. Understanding this broader evolution requires a perspective that can explain not only technological innovation, but also how heterogeneous actors become connected and stabilized over time.

To address this issue, this study adopts Actor-Network Theory (ANT) as a theoretical lens. ANT explains socio-technical development as a process of translation through which actors are identified, aligned, enrolled, and mobilized into more durable networks (Callon, 1984; Latour, 2005; Orlikowski & Scott, 2008). From this perspective, Edu-Metaverse can be understood as an evolving actor-network rather than a single technology or platform.

Empirically, this study uses MPA to trace the dominant knowledge trajectory of Edu-Metaverse research. MPA is particularly suitable for identifying the most significant citation paths within a knowledge network and reconstructing how a field develops over time (Hummon & Doreian, 1989; Liu

& Lu, 2012; Liu et al., 2019). Guided by this approach, the study addresses the following research question.

How has Edu-Metaverse evolved as a process of actor-network formation?

By combining MPA with an ANT perspective, this study makes several key contributions. It reconstructs the dominant developmental trajectory of Edu-Metaverse research, providing a structured understanding of how the field has evolved over time. It further reconceptualizes Edu-Metaverse as an evolving actor-network, highlighting how heterogeneous actors are progressively aligned into socio-technical educational systems. In addition, it identifies three core processes–alignment, mediation, and expansion–that explain how these systems emerge and stabilize. Finally, the study outlines a future research agenda to guide subsequent work on the empirical examination, governance, and contextual variation of Edu-Metaverse.

Theoretical Background: Actor-Network Formation

ANT views technological systems as outcomes of interactions among heterogeneous actors rather than as isolated artifacts (Callon, 1984; Law, 1992; Latour, 2005). In this view, agency is distributed across both human and non-human entities. In educational contexts, learners, teachers, platforms, virtual environments, algorithms, and institutional rules can all be treated as actors shaping how educational systems function (Orlikowski & Scott, 2008; Yoo et al., 2010).

A central concept in ANT is translation, the process through which networks are formed and stabilized. Callon (1984) conceptualizes this process through four moments: Problematization, where a problem is defined and key actors are positioned; Interessement, where actors are drawn into a proposed configuration; Enrolment, where roles and relationships become more structured; and Mobilization, where the network expands and gains broader stability and influence.

Applying ANT to Edu-Metaverse shifts the analytical focus from isolated technologies toward the formation of an interconnected educational environment. Rather than asking only what immersive technologies do, ANT helps explain how they become pedagogically meaningful through alignment with educators, learners, tasks, and institutions. This makes ANT especially useful for understanding the long-term development of Edu-Metaverse as a socio-technical formation. The framework is summarized in Table 1 and serves as the basis for analyzing the development trajectory of Edu-Metaverse research.

Table 1. Adapted Dimensions of Actor-Network Formation in Edu-Metaverse

Dimensions	Definition	Research focus
Actors	Human and non-human entities participating in the network.	Identify the dominant actors in each stage of Edu-Metaverse development.
Translation Goal	The educational problem or objective through which actors are aligned.	Examine what educational purposes drive actor alignment in each phase.
Network Configuration	The ways actors, technologies, and pedagogical arrangements are assembled and stabilized.	Observe how socio-technical relations are formed in each developmental phase.
Educational Outcomes	The educational consequences produced through the stabilization and expansion of actor-networks.	Assess how each phase contributes to broader educational development.

Method

This study employs MPA to reconstruct the dominant developmental trajectory of Edu-Metaverse research. MPA is a citation-based method that identifies significant knowledge flows within a scientific field by analyzing the structure of citation networks (Hummon & Doreian, 1989; Liu & Lu, 2012). It has been widely applied to trace the evolution of interdisciplinary research domains (Liu et al., 2019).

The dataset was collected from Web of Science (WoS), which provides high-quality and interdisciplinary coverage across education and technology research (Mongeon & Paul-Hus, 2016). The search strategy combined metaverse- and immersive technology-related terms (e.g., "metaverse," "virtual reality," "augmented reality," "virtual world") with education-related terms (e.g., "learning," "teaching"). The detailed query is presented in Table 2. Only articles, review articles, and early access publications indexed in SSCI and SCI-EXPANDED were included. The time span (2006–2025) captures the evolution from early virtual environments to recent metaverse-based education (John, 2007; Radianti et al., 2020). The final dataset consists of 3,049 publications.

Table 2. Search strategy and keywords used

Database	Web of science
Keywords	TS=("Metaverse" OR "Virtual World*" OR "Augmented Reality" OR "Virtual Reality" OR "Immersive VR" OR "Mixed Reality" OR "Avatars" OR "extended reality" OR "XR" OR "AR" OR "MR" OR "VR") AND ("educat*" OR "learn*" OR "E-learning" OR "student*" OR "teach*")
Type of journal article	Article or Early Access or Review
Timespan	January 1, 2006 – December 31, 2025
WoS index	Social Sciences Citation Index (SSCI) or Science Citation Index Expanded (SCI-EXPANDED)
WoS categories	Education Educational Research
Total Results	3,049 publications

A citation network was constructed in which nodes represent publications and edges represent citation relationships. MPA was conducted using the Search Path Link Count (SPLC) algorithm to evaluate the structural importance of citation links (Hummon & Doreian, 1989; Liu & Lu, 2012). This study adopts Key-Route MPA, which identifies high-weight links and expands them into representative paths, enabling the capture of the dominant development trajectory of the field (Liu et al., 2019; Xiao et al., 2014).

The identified main path publications were then analyzed qualitatively as representative knowledge nodes (Liu et al., 2019). Finally, the trajectory was interpreted through Actor-Network Theory (ANT) to explain how Edu-Metaverse evolves as a process of actor-network formation (Callon, 1984; Latour, 2005; Orlikowski & Scott, 2008).

Findings and Analysis

The Key-Route main path reveals that Edu-Metaverse developed through four broad phases that align closely with ANT's moments of translation. These phases capture how heterogeneous actors, including technologies, educators, and learners, become progressively aligned and stabilized within the emerging educational ecosystem (see Figure 1).

Phase 1: Problematization

The earliest phase centers on defining the educational problems that immersive technologies might solve. Research during this stage explored Web3D, virtual humans, and 3D simulation systems as possible responses to the limitations of conventional learning environments, especially in relation to visualization, interaction, and experiential learning (John, 2007; Karpouzis et al., 2007; Mzoughi et al., 2007). At this stage, the actor-network remained relatively simple and technology-centered. Technologies were positioned as promising solutions, while educators and learners were mainly treated as users of emerging systems.

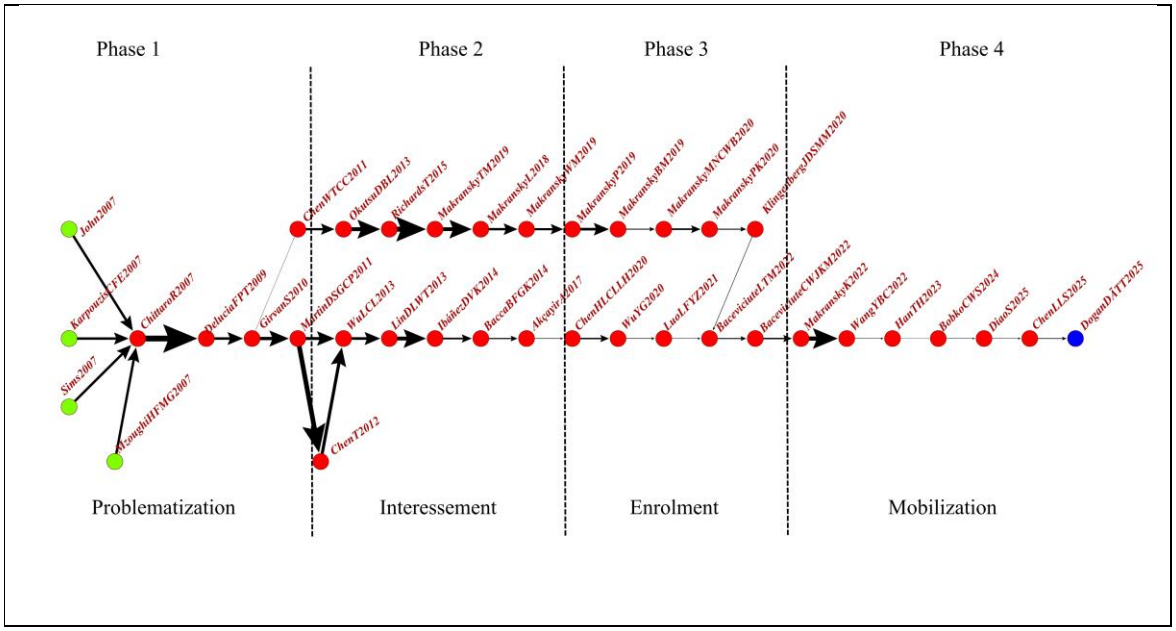


Figure 1. Key-route main paths on Edu-Metaverse

Phase 2: Interessement

The second phase reflects a stronger alignment between immersive technologies and pedagogical objectives. Research increasingly examined how augmented reality, virtual worlds, and simulation environments could support specific instructional tasks and formal learning contexts (Chen & Tsai, 2012; Lin et al., 2013; Wu et al., 2013). In ANT terms, this phase represents interessement because actors were being drawn into more stable relationships: technologies were no longer explored only as possibilities, but as tools to support teaching, collaboration, and learning design (Okutsu et al., 2013).

Phase 3: Enrolment

The third phase marks the stabilization of immersive virtual reality as a structured educational configuration. Studies moved beyond asking whether immersive environments could be useful and instead examined the mechanisms through which they influence learning outcomes, such as presence, cognitive load, self-efficacy, and instructional design quality (Makransky & Petersen, 2019; Makransky et al., 2019; Wu et al., 2020; Klingenberg et al., 2020). During this phase, the network became more internally organized: pedagogical agents, feedback systems, assessment mechanisms, and theoretically informed instructional strategies were increasingly incorporated into immersive learning environments.

Phase 4: Mobilization

The most recent phase extends Edu-Metaverse toward ecosystem-level configurations. Here, the field shifts from individual immersive experiences to interconnected platforms involving collaboration, AI-supported learning, institutional coordination, and governance concerns (Wang et al., 2022; Hwang & Chien, 2022; Han et al., 2023). In this phase, the actor-network expands significantly. Learners and

teachers are repositioned as ecosystem participants, while platforms, intelligent agents, and institutional arrangements become central to educational value creation. Edu-Metaverse is no longer framed as a discrete learning tool, but as an evolving educational infrastructure.

Discussion

The findings suggest that Edu-Metaverse evolves as a process of actor-network formation, in which educational value emerges through the progressive alignment of heterogeneous actors rather than technological advancement alone (Callon, 1984; Orlikowski & Scott, 2008). In the early stage, immersive technologies remain fragmented and produce inconsistent outcomes without strong pedagogical integration (De Lucia et al., 2009; Bacca et al., 2014). As the field develops, alignment between technological affordances and instructional design becomes critical, improving learning effectiveness in collaborative and contextual environments (Lin et al., 2013; Wu et al., 2013).

The stabilization of the network further depends on the mediation of learning processes through cognitive and affective mechanisms such as presence and cognitive load (Makransky & Petersen, 2019; Makransky et al., 2019). More recently, the network expands toward ecosystem-level configurations, where platforms, AI, and institutional actors enable scalable and collaborative learning systems (Wang et al., 2022; Han et al., 2023). As illustrated in Figure 2, this evolution can be understood through three interrelated processes: alignment, mediation, and expansion, which collectively explain the transition of Edu-Metaverse from fragmented applications to an integrated socio-technical educational infrastructure.

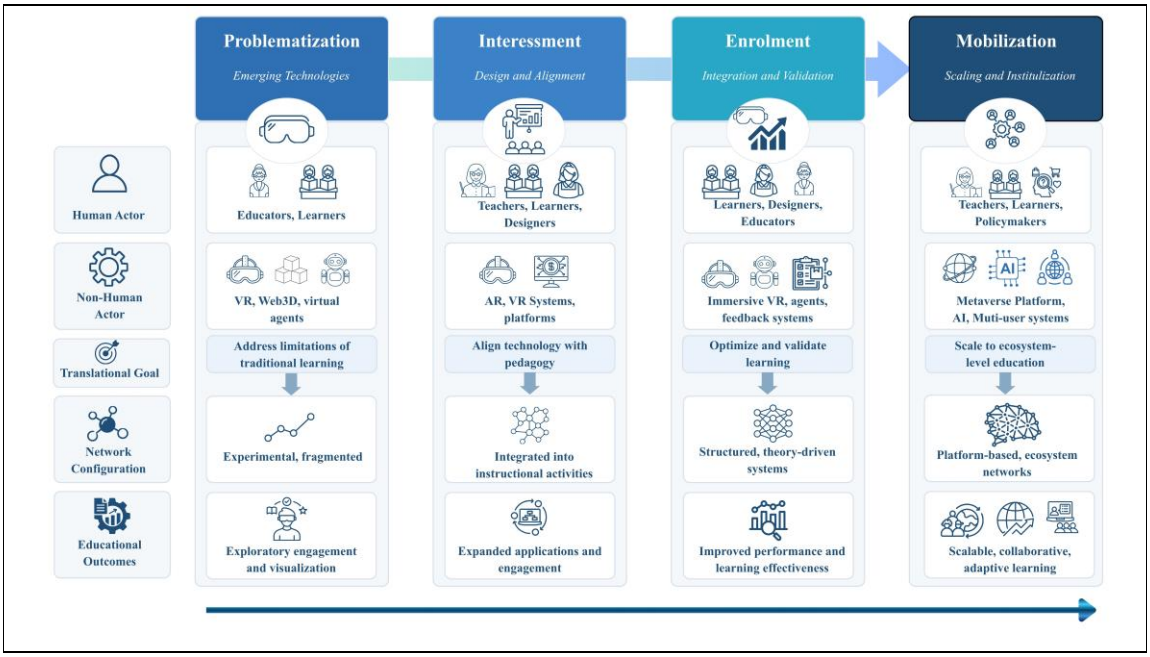


Figure 2. Edu-Metaverse Network Formation Framework

Implications and Limitations

Theoretical Contributions

This study contributes to the Edu-Metaverse literature by offering a process-oriented and socio-technical perspective on its development. By integrating Key-Route MPA with Actor-Network Theory, the study moves beyond fragmented, technology-centric views and reconstructs the dominant developmental trajectory of the field. It further advances theory by conceptualizing Edu-Metaverse as an evolving actor-network, emphasizing that its development depends on the progressive alignment of heterogeneous actors rather than technological advancement alone (Callon, 1984; Orlikowski & Scott, 2008). In addition, the study identifies three core processes–alignment, mediation, and expansion–

which provide a parsimonious framework for explaining how socio-technical educational systems emerge and stabilize over time.

Practical Contributions

The findings offer important implications for educators, institutions, and technology designers. They highlight that the effectiveness of Edu-Metaverse depends on the alignment between technology and pedagogy, rather than the adoption of advanced technologies alone (Radianti et al., 2020). The results also emphasize the role of educators as active coordinators of socio-technical systems, responsible for integrating immersive technologies into meaningful learning experiences. As Edu-Metaverse evolves toward ecosystem-level configurations, institutions should place greater emphasis on platform integration, collaboration, and governance, including issues related to accessibility, scalability, and the responsible use of AI (Wang et al., 2022; Dwivedi et al., 2022).

Limitations

This study has several limitations. The dataset is restricted to Web of Science, which may exclude relevant publications from other databases or emerging conference venues (Mongeon & Paul-Hus, 2016). While MPA effectively captures dominant knowledge trajectories, it may overlook less cited but potentially innovative research (Hummon & Doreian, 1989; Liu et al., 2019). In addition, the ANT-based interpretation involves a degree of qualitative judgment in mapping phases and actor configurations. Finally, the study focuses on the structural evolution of the research field rather than empirical validation in specific educational contexts. Future research could combine bibliometric analysis with empirical methods to further examine how Edu-Metaverse actor-networks operate in practice.

Future Research Agenda

This study opens several avenues for future research on Edu-Metaverse as an evolving actor-network. Future studies could move beyond bibliometric analysis to empirically examine how actor-network formation unfolds in practice, using qualitative or mixed-method approaches such as case studies, experiments, or longitudinal designs (Orlikowski & Scott, 2008; Yin, 2018). Such efforts would help validate the processes of alignment, mediation, and expansion identified in this study within real educational contexts.

Greater attention is also needed to the micro-level mechanisms through which actors become aligned. Examining the roles of educators, learners, and intelligent systems in shaping learning experiences can provide deeper insights into how pedagogical practices, user behaviors, and technological affordances interact over time (Yoo et al., 2010; Makransky & Petersen, 2019).

As Edu-Metaverse continues to evolve toward ecosystem-level configurations, further research should address broader issues of governance, ethics, and institutional dynamics. Topics such as accessibility, data privacy, platform dependency, and the responsible use of artificial intelligence are critical for ensuring sustainable and inclusive development (Dwivedi et al., 2022; Wang et al., 2022).

In addition, comparative and cross-context studies would enrich understanding of how Edu-Metaverse actor-networks differ across educational levels, cultural settings, and technological infrastructures. Such perspectives can help clarify the conditions under which socio-technical educational systems emerge, stabilize, and scale in diverse contexts (Radianti et al., 2020; Won et al., 2023).

Conclusion

This study examines the development of Edu-Metaverse through the lens of ANT and Key-Route MPA. By reconstructing the dominant knowledge trajectory of 3,049 publications, the findings show that Edu-Metaverse evolves as a process of actor-network formation rather than as a sequence of isolated technological advancements. Specifically, the field progresses through four phases–Problematization, Interessement, Enrolment, and Mobilization–through which technologies, educators, learners, and

institutional arrangements become progressively aligned. The results highlight that the emergence of Edu-Metaverse is driven not only by technological innovation but by the alignment between technological capabilities and pedagogical practices. Over time, this alignment enables the transition from experimental applications to more stable and scalable educational systems, and ultimately toward ecosystem-level configurations. Overall, this study contributes to a process-oriented perspective on Edu-Metaverse development, offering a clearer theoretical explanation of how socio-technical educational systems emerge and evolve. It also provides a foundation for future research to further examine the dynamics of actor alignment and system-level integration in next-generation digital learning environments.

References

Bacca, J., Baldiris, S., Fabregat, R., Graf, S., and Kinshuk. 2014. "Augmented Reality Trends in Education: A Systematic Review of Research and Applications," *Educational Technology & Society* (17:4), pp. 133–149. https://www.jstor.org/stable/jeductechsoci.17.4.133

Callon, M. 1984. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay," *The Sociological Review* (32:S1), pp. 196–233. https://doi.org/10.1111/j.1467-954X.1984.tb00113.x

Chen, C. M., and Tsai, Y. N. 2012. "Interactive Augmented Reality System for Enhancing Library Instruction in Elementary Schools," *Computers & Education* (59:2), pp. 638–652. https://doi.org/10.1016/j.compedu.2012.03.001

De Lucia, A., Francese, R., Passero, I., and Tortora, G. 2009. "Development and Evaluation of a Virtual Campus on Second Life: The Case of SecondDMI," *Computers & Education* (52:1), pp. 220–233. https://doi.org/10.1016/j.compedu.2008.08.001

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., et al. 2022. "Metaverse Beyond the Hype: Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy," *International Journal of Information Management* (66), 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Han, Z., Tu, Y., and Huang, C. 2023. "A Framework for Constructing a Technology-Enhanced Education Metaverse: Learner Engagement with Human–Machine Collaboration," *IEEE Transactions on Learning Technologies* (16:6), pp. 1179–1189. https://doi.org/10.1109/TLT.2023.3257511

Hummon, N. P., and Doreian, P. 1989. "Connectivity in a Citation Network: The Development of DNA Theory," *Social Networks* (11:1), pp. 39–63. https://doi.org/10.1016/0378-8733(89)90017-8

Hwang, G.-J., and Chien, S.-Y. 2022. "Definition, Roles, and Potential Research Issues of the Metaverse in Education: An Artificial Intelligence Perspective," *Computers and Education: Artificial Intelligence* (3), 100082. https://doi.org/10.1016/j.caeai.2022.100082

John, N. W. 2007. "The Impact of Web3D Technologies on Medical Education and Training," *Computers & Education* (49:1), pp. 19–31. https://doi.org/10.1016/j.compedu.2005.06.003

Karpouzis, K., Caridakis, G., Fotinea, S.-E., and Efthimiou, E. 2007. "Educational Resources and Implementation of a Greek Sign Language Synthesis Architecture," *Computers & Education* (49:1), pp. 54–74. https://doi.org/10.1016/j.compedu.2005.06.004

Klingenberg, S., Jørgensen, M. L. M., Dandanell, G., Skriver, K., Mottelson, A., and Makransky, G. 2020. "Investigating the Effect of Teaching as a Generative Learning Strategy When Learning Through Desktop and Immersive VR," *British Journal of Educational Technology* (51:6), pp. 2115–2138. https://doi.org/10.1111/bjet.13029

Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780199256044.001.0001

Law, J. 1992. "Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity," *Systems Practice* (5:4), pp. 379–393. https://doi.org/10.1007/BF01059830

Lin, T. J., Duh, H. B. L., Li, N., Wang, H. Y., and Tsai, C. C. 2013. "An Investigation of Learners' Collaborative Knowledge Construction Performances and Behavior Patterns in an Augmented Reality Simulation System," *Computers & Education* (68), pp. 314–321. https://doi.org/10.1016/j.compedu.2013.05.011

Liu, J. S., and Lu, L. Y. Y. 2012. "An Integrated Approach for Main Path Analysis: Development of the Hirsch Index as an Example," *Journal of the American Society for Information Science and Technology* (63:3), pp. 528–542. https://doi.org/10.1002/asi.21692

Liu, J. S., Lu, L. Y. Y., and Ho, M. H. C. 2019. "A Few Notes on Main Path Analysis," *Scientometrics* (119:1), pp. 379–391. https://doi.org/10.1007/s11192-019-03034-x

Makransky, G., and Petersen, G. B. 2019. "Investigating the Process of Learning with Desktop Virtual Reality: A Structural Equation Modeling Approach," *Computers & Education* (134), pp. 15–30. https://doi.org/10.1016/j.compedu.2019.02.002

Makransky, G., Terkildsen, T. S., and Mayer, R. E. 2019. "Adding Immersive Virtual Reality to a Science Lab Simulation Causes More Presence but Less Learning," *Learning and Instruction* (60), pp. 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Mongeon, P., and Paul-Hus, A. 2016. "The Journal Coverage of Web of Science and Scopus: A Comparative Analysis," *Scientometrics* (106:1), pp. 213–228. https://doi.org/10.1007/s11192-015-1765-5

Mzoughi, T., Herring, S. D., Foley, J. T., Morris, M. J., and Gilbert, P. J. 2007. "WebTOP: A 3D Interactive System for Teaching and Learning Optics," *Computers & Education* (49:1), pp. 110–129. https://doi.org/10.1016/j.compedu.2005.06.008

Okutsu, M., DeLaurentis, D., Brophy, S., and Lambert, J. 2013. "Teaching an Aerospace Engineering Design Course via Virtual Worlds: A Comparative Assessment of Learning Outcomes," *Computers & Education* (60:1), pp. 288–298. https://doi.org/10.1016/j.compedu.2012.07.012

Orlikowski, W. J., and Scott, S. V. 2008. "Sociomateriality: Challenging the Separation of Technology, Work and Organization," *Academy of Management Annals* (2:1), pp. 433–474. https://doi.org/10.1080/19416520802211644

Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. 2020. "A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda," *Computers & Education* (147), 103778. https://doi.org/10.1016/j.compedu.2019.103778

Sripan, T., and Jeerapattanatorn, P. 2025. "Metaverse-Based Learning: A Comprehensive Review of Current Trends, Challenges, and Future Implications," *Contemporary Educational Technology* (17:3), Article ep584. https://doi.org/10.30935/cedtech/16434

Wang, X., Young, G. W., Plechatá, A., McGuckin, C., and Makransky, G. 2023. "Utilizing Virtual Reality to Assist Social Competence Education and Social Support for Children from Under-Represented Backgrounds," *Computers & Education* (200), Article 104815. https://doi.org/10.1016/j.compedu.2023.104815

Won, M., Ungu, D. A. K., Matovu, H., Treagust, D. F., Tsai, C.-C., Park, J., Mocerino, M., and Tasker, R. 2023. "Diverse Approaches to Learning with Immersive Virtual Reality Identified from a Systematic Review," *Computers & Education* (195), Article 104701. https://doi.org/10.1016/j.compedu.2022.104701

Wu, B., Yu, X., and Gu, X. 2020. "Effectiveness of Immersive Virtual Reality Using Head-Mounted Displays on Learning Performance: A Meta-Analysis," *British Journal of Educational Technology* (51:6), pp. 1991–2005. https://doi.org/10.1111/bjet.13023

Wu, H. K., Lee, S. W. Y., Chang, H. Y., and Liang, J. C. 2013. "Current Status, Opportunities and Challenges of Augmented Reality in Education," *Computers & Education* (62), pp. 41–49. https://doi.org/10.1016/j.compedu.2012.10.02

Wu, T., and Hao, F. 2023. "Edu-Metaverse: Concept, Architecture, and Applications," *Interactive Learning Environments* (32:8), pp. 4352–4379. https://doi.org/10.1080/10494820.2023.2198567

Xiao, Y., Lu, L. Y. Y., Liu, J. S., and Zhou, Z. 2014. "Knowledge Diffusion Path Analysis of Data Quality Literature: A Main Path Analysis," *Journal of Informetrics* (8:3), pp. 594–605. https://doi.org/10.1016/j.joi.2014.05.001

Yin, R. K. 2018. *Case Study Research and Applications: Design and Methods* (6th ed.). Thousand Oaks, CA: Sage Publications.

Yoo, Y., Henfridsson, O., and Lyytinen, K. 2010. "The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research," *Information Systems Research* (21:4), pp. 724–735. https://doi.org/10.1287/isre.1100.0322

B4

Generation Z Generative AI Travel Decision-Making Value Co-Creation

Ganghao Zhang, Xinyuan Xu, Mengqiang Ma, Yuxin Li, Jingjing Liu, Fuying Xu,
Tongzheng Wang, Safa Ait Larbi (Shandong Technology and Business University),
Han Heesup (Sejong University)

Focusing on Generation Z, this study combines SEM and fsQCA to explore factors influencing their sustained AI adoption in travel decision-making from a human-machine value co-creation perspective. Results show performance expectancy, hedonic motivation, and habit positively affect value co-creation; perceived risk and individual innovativeness moderate outcomes. Four configurations for high continued usage intention are identified, with value co-creation playing a mediating role. This study reveals a new human-machine collaboration paradigm and provides theoretical practical support for smart tourism.

KEY WORDS

Generation Z, generative AI, travel decision-making, value co-creation behavior, continued use intention

B4

Extending Perceived Risk Framework in Live Streaming E-Commerce: An Information Ecology Perspective

Zitong Li, Zhaohong Liang and Xiaosong Wu (Beijing Normal Univ.)

Live streaming commerce (LSC) differs from traditional online shopping in its real-time, highly interactive nature-features that can themselves become sources of consumer risk. However, existing perceived risk frameworks, developed for conventional e commerce context, fail to take such emerging dynamics into account. Drawing on information ecology theory, we develop an extending framework in which the alienations of the interactions among information, information persons, and information environment are conceptualized as risk sources. Through a theoretical review of 53 core articles using both manual and AI-assisted dual coding, we identify four risks: cognitive overload risk, technological mediation risk, social interaction risk, and market distortion risk. These findings extend perceived risk theory by shifting focus from outcome-based to process-based risks.

KEY WORDS

live streaming commerce, consumer perceived risk, information ecology theory, risk dimensions, theoretical review

Extending Perceived Risk Framework in Live Streaming E-Commerce: An Information Ecology Perspective

Li, Z.[a], Liang, Z.[b] and Wu, X.[c]
[a] Bay Area International Business School, Beijing Normal University
E-mail: 202411109062@mail.bnu.edu.cn
[b] Bay Area International Business School, Beijing Normal University
E-mail: liang_zhaohong@163.com
[c] Bay Area International Business School, Beijing Normal University
E-mail: wuxiaosong@bnu.edu.cn

Abstract

Live streaming commerce (LSC) differs from traditional online shopping in its real-time, highly interactive nature—features that can themselves become sources of consumer risk. However, existing perceived risk frameworks, developed for conventional e-commerce context, fail to take such emerging dynamics into account. Drawing on information ecology theory, we develop an extending framework in which the alienations of the interactions among information, information persons, and information environment are conceptualized as risk sources. Through a theoretical review of 53 core articles using both manual and AI-assisted dual coding, we identify four risks: cognitive overload risk, technological mediation risk, social interaction risk, and market distortion risk. These findings extend perceived risk theory by shifting focus from outcome-based to process-based risks.

Keywords: live streaming commerce, consumer perceived risk, information ecology theory, risk dimensions, theoretical review

Introduction

Live streaming commerce (LSC) combines live video, social interaction, and instant purchasing (Zhang and Tang 2023), creating a highly engaging shopping environment that has rapidly expanded within e-commerce. Compared with traditional online shopping, consumers in LSC often face higher levels of perceived risk, largely due to its real-time and interactive nature. Understanding how these risks emerge and influence consumer decision-making is therefore both important and challenging.

Most existing studies on LSC take an enabling perspective, focusing on how streamers build trust and reduce uncertainty to encourage purchases (Lu and Chen 2021; Wen et al. 2024). In this line of research, risk is usually treated as something to overcome, rather than something to be systematically examined. The classical six-dimensional risk framework (Jacoby and Kaplan 1972; Stone and Grønhaug 1993), while widely used in traditional e-commerce, does not fully capture the unique characteristics of LSC. Features such as real-time broadcasting (Wongkitrungrueng and Assarut 2020), synchronous interaction (Joo and Yang 2023), and personalized service (Sun et al. 2019) not only reshape the shopping experience but may also introduce new types of risk. Although some recent studies have started to explore these emerging risks (Chen and Yang 2023), the findings are still scattered and lack a coherent framework.

To address this gap, this study adopts information ecology theory (Nardi and O'Day 1999), which views risk as arising from disruptions in the interactions among information, information users, and the

surrounding environment. In the context of LSC, consumer risk can therefore be understood as the result of how these interactions are organized and influenced by technological, commercial, and social factors. This perspective shifts the focus from simply categorizing risks to explaining how they are generated, leading to four key dimensions: cognitive overload, technological mediation, social interaction, and market distortion.

Based on this approach, this study conducts a systematic literature review with two main objectives. First, it aims to develop an integrated framework of consumer perceived risk in LSC grounded in information ecology theory. Second, it seeks to synthesize existing research on risk mitigation while clarifying how different types of risk may shape consumer decision-making in live streaming commerce.

Literature Review and Theoretical Framework

Consumer Perceived Risk: A Classical Framework

The concept of perceived risk, introduced by Bauer (1960) and defined as the expectation of adverse outcomes under uncertainty (Stone and Grønhaug 1993), has been central to consumer decision-making research. Jacoby and Kaplan (1972) identified six dimensions—financial, performance, psychological, social, time, and physical—that shape overall risk perception. This framework has been widely applied in traditional retail and early e-commerce settings (Featherman and Pavlou 2003; Pavlou 2003).

Risk Research in Live Streaming Commerce

Existing LSC research primarily adopts an enabling perspective, focusing on how streamers build trust and reduce uncertainty (Lu and Chen 2021; Wen et al. 2024). Risk is typically treated as a barrier, leaving systematic investigation limited. Although the classical six-dimensional framework remains robust in traditional e-commerce, it has limitations in the LSC context. Some studies have identified emerging risks (Chen and Yang 2023; Qu et al. 2023), but these insights remain fragmented. What is missing is a unifying lens that explains not only what these risks are, but also how they arise from LSC's unique dynamics.

Information Ecology Theory: A New Analytical Lens

To address this issue, this study draws on information ecology theory (IET) (Nardi and O'Day 1999). IET views an information system as consisting of three core elements: information, information persons, and the information environment, and suggests that risks may arise when interactions among these elements become misaligned or disrupted. In this framework, individuals create and use information, while the environment shapes how information flows and how actions are taken. When these interactions break down—due to factors such as information overload, technological constraints, or social influences—various forms of risk may emerge.

Applying this perspective to LSC helps clarify how these elements work in this specific context. Information includes product details, pricing, promotional messages, and real-time comments from viewers. Information users include both streamers, who present and organize the information, and consumers, who interpret and respond to it. The information environment consists of the platform itself, its interaction rules, and the social dynamics among users. In this setting, risks do not only come from the products themselves, but also from how information is communicated and how users interact within the platform.

Theoretical Development: A Four-Dimensional Risk Framework

Building on this view, we decompose the information environment into three sub-environments—technological, commercial, and social. The technological environment includes algorithms, interface design, and system reliability; the commercial environment covers promotional rules, pricing strategies, and competition; and the social environment encompasses streamer–follower relationships, peer dynamics, and collective sentiment. This lens shifts the focus from outcome-based risks (e.g., financial loss) to process-based risks arising from interaction dynamics. We propose a four-dimensional

classification in which the core interaction between information and information persons becomes alienated either directly—through overload or contradiction—or through technological, social, or commercial mediation. To assess whether these dimensions appear in prior research, we conducted a theoretical review; the following sections present the methodology and findings.

First, direct interaction disrupted by information overload or contradiction gives rise to cognitive overload risk, leading to decision fatigue or premature choices (Joo and Yang 2023; Ma 2021a). Second, mediation by the technological environment—through unreliability, opacity, or manipulation—generates technological mediation risk, including algorithmic bias, system failures, data privacy concerns, and nudging interface designs (Chen et al. 2024; Pavlou 2003). Third, embedding within the social environment—where social bonds may be strategically exploited—leads to social interaction risk, including parasocial exploitation, peer pressure, and emotional manipulation (Liao et al. 2023; Rungruangjit 2022). Fourth, shaping by the commercial environment—where market rules distort transparency—produces market distortion risk, manifesting in deceptive pricing, complex promotions, and unclear accountability (Chen and Zhang 2023; Guo et al. 2021). These dimensions often co-occur and reinforce one another, reflecting the systemic nature of risk in LSC. Figure 1 illustrates the framework.

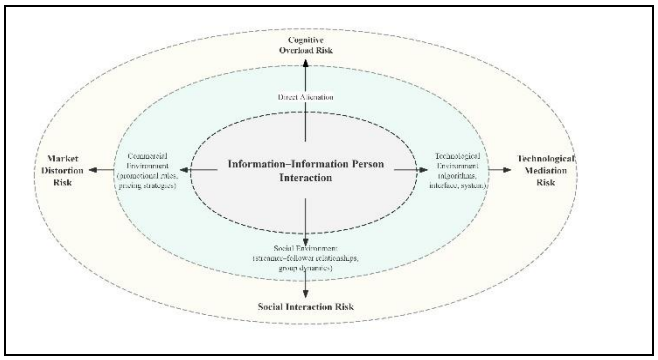


Figure 1. Four-Dimensional Risk Framework

Methodology

Research Approach

This study adopts a theoretical review approach (Paré et al. 2015), which synthesizes existing research into a higher-order conceptual framework. Theoretical reviews typically begin with a broad question refined as evidence accumulates and employ a systematic approach to ensure comprehensive coverage of the literature (Paré et al. 2015; Webster and Watson 2002).

Literature Search Strategy

We conducted a systematic literature search following PRISMA guidelines (Table 1). The search covered 2015–2025 and returned 517 records, leaving 337 after deduplication.

Table 1. Literature Search Strategy

Database	Timeframe	Initial Results
Web of Science	2015–2025	215
Scopus	2015–2025	302
Total after deduplication		337

Study Selection Process

We screened titles and abstracts, then assessed full texts against four criteria: perceived risk focus in LSC; peer-reviewed journals or top-tier conferences; English language; and empirical or conceptual

nature. Articles meeting all criteria formed the core literature (n=53); additional articles were kept as supplementary material. Figure 2 presents the PRISMA flow diagram.

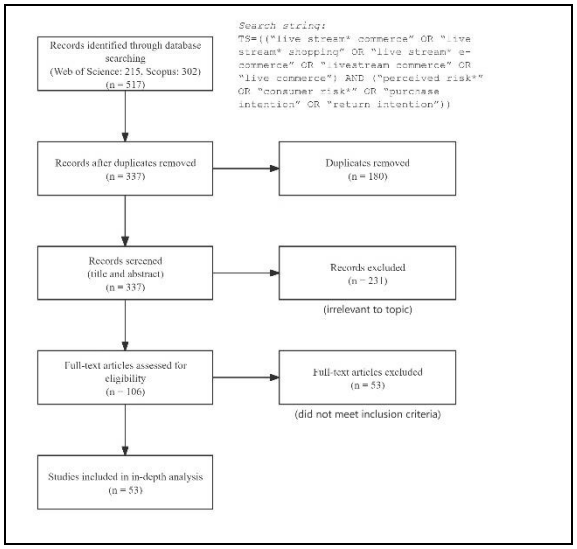


Figure 2. PRISMA Flow Diagram

Dual Coding Process

We adopted a dual coding strategy that combined manual analysis with AI-assisted support. First, the first author manually reviewed all 53 core articles and extracted risk-related passages, which were systematically recorded in a coding sheet. To complement this process, we then used DeepSeek with a structured prompt to (a) map the extracted content onto the classical six-dimensional framework and (b) identify potential new risk dimensions. The prompt was iteratively refined using three pilot articles before being applied to the full dataset. Table 2 presents an illustrative example of the coding output.

Table 2. Example of Coding Output

Article ID	A. Mapped Classical Dimensions	B. Potential New Risk Dimension	C. Analysis Notes
Ma et al. (2022)	Functional Risk: "lack of confidence in product quality"	Decision Flow Disruption Risk	Functional risk mapping "high"

Coding Reliability

To assess reliability, we conducted a macro-level consistency check. An independent researcher reviewed 11 randomly selected articles and provided brief summaries of the core risks identified in each study. These summaries were then compared with the risk dimensions generated through AI-assisted coding, with a focus on whether they aligned at a broader categorical level (e.g., technological, social, or emotional risks). Table 3 presents several representative examples.

Overall, 9 of the 11 articles (82%) showed a high level of consistency. The remaining two cases (Joo and Yang 2023; Liu and Wang 2023) were considered moderately consistent, but still aligned with the broader risk categories. Taken together, these findings suggest that the AI-assisted coding approach is able to identify risk-related domains in a reliable manner.

Table 3. Macro Consistency Check (11 Articles)

Article ID	Reviewer's Summary	AI-Identified Risk	Consistency
Pavlou (2007)	Information asymmetry, seller opportunism, privacy concerns	Principal-Agent Relationship Alienation Risk	High
Qu et al. (2023)	Scarcity-induced impulse buying, emotions	Cognitive Narrowing Risk from Immersion	High

Gao et al. (2023)	Virtual streamer characteristics, social presence, technology + social	Virtual Influencer Authenticity Risk	High

Theory Integration Method

We adopted a three-step approach grounded in information ecology theory. First, we identified the source of each AI-generated risk, distinguishing among information, technological, social, and commercial factors. Next, each risk was mapped to one of the four core interactions. Finally, the risks were assigned to the corresponding overarching dimension.

Through this process, a wide range of previously dispersed risk concepts was integrated into four higher-level dimensions. Table 4 presents the integration results based on the 11 sample articles.

Table 4. Integration Results for 11 sample Articles

Article ID	AI-Identified Risk	Final Dimension
Pavlou (2007)	Principal-Agent Relationship Alienation Risk	Market Distortion Risk
Dong et al. (2022)	Values Exploitation Risk	Social Interaction Risk
Zhang et al. (2024)	Content Strategy Misalignment Risk	Market Distortion Risk
Yan et al. (2023)	Content Strategy Misalignment Risk	Market Distortion Risk
Qu et al. (2023)	Cognitive Narrowing Risk from Immersion	Cognitive Overload Risk
Luo et al. (2024)	Collective Irrationality Risk from Emotional Mobilization	Social Interaction Risk
Liu and Wang (2023)	Cooperation Sustainability Risk	Market Distortion Risk
Liao et al. (2023)	Parasocial Relationship Authenticity Risk	Social Interaction Risk
Joo and Yang (2023)	Cognitive Resource Overload Risk	Cognitive Overload Risk
Huang et al. (2024)	Hedonic Value Addiction Risk	Social Interaction Risk
Gao et al. (2023)	Virtual Influencer Authenticity Risk	Technological Mediation Risk

Findings

This section presents the four risk dimensions that emerged from the theoretical review and the three-step integration method described in Section 3. Each dimension is defined within the information ecology framework, supported by representative literature, and illustrated with typical manifestations in live streaming commerce.

Cognitive Overload Risk

Cognitive overload risk refers to the risk that the volume and speed of real-time information in LSC exceed consumers' processing capacity, leading to decision fatigue or less deliberate choices. From an information ecology perspective, this reflects how disruptions in the direct interaction between information and the information person—such as excessive volume or conflicting signals—can overwhelm processing capacity and lead to cognitive overload. In such settings, consumers must process multiple streams simultaneously; when these streams become excessive or inconsistent, cognitive resources are quickly depleted.

Prior research suggests that simultaneous information flows can overwhelm viewers, leaving only a superficial sense of authenticity (Ma 2021b), while high interactivity further reduces the capacity for careful evaluation (Joo and Yang 2023). For instance, rapidly scrolling danmaku may obscure key product details, making it difficult to process key details and increasing reliance on rapid, and sometimes suboptimal, decisions.

Technological Mediation Risk

Technological mediation risk refers to the risk that platform infrastructures, algorithms, or interface designs in LSC operate in ways that consumers cannot fully observe or control, leading to reduced trust, perceived loss of control, or disrupted transactions. From an information ecology perspective, this

reflects how reliance on opaque or unreliable technological mediation—such as algorithms or interface design—can distort the interaction between consumers and information. In practice, consumers rely on system reliability and algorithmic curation to access and interpret information; when these mechanisms become opaque or unreliable—due to bias, system failures, or nudging interface designs—the interaction becomes distorted.

Prior research suggests that the anthropomorphism of virtual streamers may weaken users' sense of control (Chen et al. 2024), while privacy and security concerns remain persistent sources of uncertainty (Pavlou et al. 2007). For example, during high-traffic flash sales, system overload may freeze payment interfaces, leaving transactions incomplete or inconsistent, prompting hesitation or withdrawal, and thereby increasing decision uncertainty and reducing perceived control over purchase outcomes.

Social Interaction Risk

Social interaction risk arises when consumers in LSC's highly interactive environment face emotional manipulation, relational exploitation, or peer pressure arising from streamer-led strategies and audience interactions, leading to less deliberate purchasing decisions or reduced autonomy. From an information ecology perspective, this reflects how social relationships and interactions can reshape and potentially distort how information is interpreted and acted upon. In such settings, information is conveyed through relationships—between streamers and followers or among viewers—where emotional connections can shape how messages are interpreted; when these connections are strategically leveraged for commercial purposes, the interaction becomes distorted.

Prior research points to both parasocial bonding and collective pressure as key mechanisms. Streamers often build intimacy through personal storytelling, framing purchase requests as acts of support (Rungruangjit 2022), which may leave viewers susceptible when such relationships are perceived as strategically constructed (Liao et al. 2023). For instance, when streamers invoke personal narratives while urging immediate purchases, viewers may comply despite limited prior intention or need, leading to unintended or impulsive purchases.

Market Distortion Risk

Market distortion risk refers to the risk that promotional mechanisms and competitive dynamics in LSC shape information in ways that obscure true costs or terms, leading to misinformed or suboptimal purchasing decisions. From an information ecology perspective, this reflects how commercial incentives and market structures can shape and distort how information is presented and evaluated. In such settings, pricing strategies, promotional rules, and competitive positioning shape how information is presented; when these mechanisms become distortive—through fictitious reference prices or overly complex conditions—the interaction is compromised.

Prior research suggests that low prices may also signal inferior quality, increasing perceived uncertainty (Guo et al. 2021), while complex promotional rules can function as strategic tools that obscure true value. For example, "lowest price ever" claims often rely on inflated reference prices or restrictive eligibility conditions (Chen and Zhang 2023), making it difficult for consumers to accurately assess the actual cost and leading to decisions they might not otherwise consider, including suboptimal or impulsive purchases.

Discussion

Summary of Findings

This study extends the classical perceived risk framework by adopting an information ecology lens to explain how risk emerges in live streaming commerce. Specifically, it identifies four process-based risk dimensions—cognitive overload, technological mediation, social interaction, and market distortion—rooted in the alienation of core interactions. These process-based risks shape consumer decision-making by influencing how information is interpreted, evaluated, and acted upon in dynamic consumption contexts.

Theoretical Contributions and Practical Implications

This study contributes to three streams of research.

First, it extends the classical six-dimensional framework (Jacoby and Kaplan 1972; Stone and Grønhaug 1993) by shifting the focus from outcome-based risks (e.g., financial loss, performance failure) to process-based risks rooted in interaction alienation. Rather than viewing risk solely as a consequence of transaction uncertainty, our findings suggest that the structure of interaction itself—whether direct or mediated by technological, social, or commercial environments—can generate risk.

Second, it advances information ecology theory (Nardi and O'Day 1999) in the context of LSC by specifying the roles of information, information persons, and the information environment, and by further decomposing the environment into technological, commercial, and social domains.

Third, it identifies four interrelated risk dimensions that extend existing understandings of consumer risk in LSC. These dimensions are not independent; they often co-occur and reinforce one another, reflecting the systemic nature of risk in this context.

These findings also offer practical implications. Platforms should design interfaces that reduce information overload and improve the transparency of algorithmic processes. Streamers should clearly distinguish between personal engagement and promotional intent to avoid exploiting relational dynamics. Brands should adopt transparent pricing strategies and simplify promotional rules to reduce information distortion and consumer confusion. Regulators may also need to address process-based risks such as interface manipulation and algorithmic opacity. Importantly, the proposed framework can be operationalized in future research by translating each risk dimension into measurable constructs (e.g., perceived information load, system transparency, social pressure, and pricing complexity), which can be examined using surveys, experiments, or behavioral data.

Limitations and Future Research

Several limitations suggest directions for future research. First, as a theoretical framework, this study calls for empirical validation; future work may develop and test measurement instruments to assess its predictive value. Second, while AI-assisted coding improves efficiency, it may introduce certain biases; future studies could employ larger human coding teams or complementary methods such as experiments or structural modeling. Third, the literature sample may not fully capture cross-cultural or emerging format differences; comparative and longitudinal research would help establish boundary conditions. Finally, integrating this framework with related perspectives, such as technostress or social influence, may further enrich understanding.

Conclusion

This study extends the classical perceived risk framework by integrating information ecology theory to explain how consumer risk emerges in live streaming commerce. Drawing on a theoretical review of 53 articles, we identify four dimensions—cognitive overload, technological mediation, social interaction, and market distortion—rooted in the alienation of core interactions within the LSC ecosystem. By shifting the focus from outcome-based to process-based risks, the framework offers a structured lens for understanding and managing risk in this evolving retail context.

References

Bauer, R. A. 1960. "Consumer Behavior as Risk Taking," in R. S. Hancock (ed.), Dynamic Marketing for a Changing World, Chicago: American Marketing Association, pp. 389–398.

Chen, C., and Zhang, D. 2023. "Understanding Consumers' Live-Streaming Shopping from a Benefit–Risk Perspective," Journal of Services Marketing (37:8), pp. 973–988.

Chen, H., Shao, B., Yang, X., Kang, W., and Fan, W. 2024. "Avatars in Live Streaming Commerce: The Influence of Anthropomorphism on Consumers' Willingness to Accept Virtual Live Streamers," Computers in Human Behavior (156), p. 108216.

Chen, N., and Yang, Y. 2023. "The Role of Influencers in Live Streaming E-Commerce: Influencer Trust, Attachment, and Consumer Purchase Intention," Journal of Theoretical and Applied Electronic Commerce Research (18:3), pp. 1601–1618.

Featherman, M. S., and Pavlou, P. A. 2003. "Predicting E-Services Adoption: A Perceived Risk Facets Perspective," International Journal of Human-Computer Studies (59:4), pp. 451–474.

Guo, J., Li, Y., Xu, Y., and Zeng, K. 2021. "How Live Streaming Features Impact Consumers' Purchase Intention in the Context of Cross-Border E-Commerce? A Research Based on SOR Theory," Frontiers in Psychology (12), p. 767876.

Jacoby, J., and Kaplan, L. B. 1972. "The Components of Perceived Risk," in Proceedings of the 3rd Annual Conference of the Association for Consumer Research, M. Venkatesan (ed.), College Park, MD: Association for Consumer Research, pp. 382–392.

Joo, E., and Yang, J. 2023. "How Perceived Interactivity Affects Consumers' Shopping Intentions in Live Stream Commerce: Roles of Immersion, User Gratification and Product Involvement," Journal of Research in Interactive Marketing (17:5), pp. 754–772.

Liao, J., Chen, K., Qi, J., Li, J., and Yu, I. Y. 2023. "Creating Immersive and Parasocial Live Shopping Experience for Viewers: The Role of Streamers' Interactional Communication Style," Journal of Research in Interactive Marketing (17:1), pp. 140–155.

Liu, B., and Wang, W. 2023. "Live Commerce Retailing with Online Influencers: Two Business Models," International Journal of Production Economics (255), p. 108715.

Lu, B., and Chen, Z. 2021. "Live Streaming Commerce and Consumers' Purchase Intention: An Uncertainty Reduction Perspective," Information & Management (58:7), p. 103509.

Ma, Y. 2021a. "Elucidating Determinants of Customer Satisfaction with Live-Stream Shopping: An Extension of the Information Systems Success Model," Telematics and Informatics (65), p. 101707.

Ma, Y. 2021b. "To Shop or Not: Understanding Chinese Consumers' Live-Stream Shopping Intentions from the Perspectives of Uses and Gratifications, Perceived Network Size, Perceptions of Digital Celebrities, and Shopping Orientations," Telematics and Informatics (59), p. 101562.

Nardi, B. A., and O'Day, V. 1999. Information Ecologies: Using Technology with Heart, Cambridge, MA: MIT Press.

Paré, G., Trudel, M.-C., Jaana, M., and Kitsiou, S. 2015. "Synthesizing Information Systems Knowledge: A Typology of Literature Reviews," Information & Management (52:2), pp. 183–199.

Pavlou, P. 2003. "Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model," International Journal of Electronic Commerce (7:3), pp. 101–134.

Pavlou, P., Liang, H., and Xue, Y. 2007. "Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective," MIS Quarterly (31), pp. 105–136.

Qu, Y., Khan, J., Su, Y., Tong, J., and Zhao, S. 2023. "Impulse Buying Tendency in Live-Stream Commerce: The Role of Viewing Frequency and Anticipated Emotions Influencing Scarcity-Induced Purchase Decision," Journal of Retailing and Consumer Services (75), p. 103534.

Rungruangjit, W. 2022. "What Drives Taobao Live Streaming Commerce? The Role of Parasocial Relationships, Congruence and Source Credibility in Chinese Consumers' Purchase Intentions," Heliyon (8:6), p. e09676.

Stone, R. N., and Grønhaug, K. 1993. "Perceived Risk: Further Considerations for the Marketing Discipline," European Journal of Marketing (27:3), pp. 39–50.

Sun, Y., Shao, X., Li, X., Guo, Y., and Nie, K. 2019. "How Live Streaming Influences Purchase Intentions in Social Commerce: An IT Affordance Perspective," Electronic Commerce Research and Applications (37), p. 100886.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," MIS Quarterly (26:2), pp. xiii–xxiii.

Wen, L., Ma, S., and Lyu, S. 2024. "The Influence of Internet Celebrity Anchors' Reputation on Consumers' Purchase Intention in the Context of Digital Economy: From the Perspective of Consumers' Initial Trust," Applied Economics (56:60), pp. 9189–9210.

Wongkitrungrueng, A., and Assarut, N. 2020. "The Role of Live Streaming in Building Consumer Trust and Engagement with Social Commerce Sellers," Journal of Business Research (117), pp. 543–556.

Zhang, T., and Tang, Z. 2023. "Should Manufacturers Open Live Streaming Shopping Channels?" Journal of Retailing and Consumer Services (71), p. 103229.

C 트랙

C1 Responsible AI & Digital Governance / Session Chair : Wenting Liu (Singapore Univ. of Social Sciences) 09:00~10:20

- **Exploring Ethical Challenges of Artificial Intelligence in Commercial Contexts**
  Ling-Pu Huang and Chen-Hao Huang (National Cheng Kung Univ.)
- **How AI Transparency Affects AI Appreciation: The Mediating Roles of Trusting Beliefs**
  Hua Liu (Anhui Univ.), Xiaofan Tang (Anhui Medical Univ.) and Jinmei Yin (Nanjing Univ. of Aeronautics and Astronautics)
- **AI-Driven Intelligent Policy Monitoring Method: Measuring the Intensity and Synergy of China's Data Element Market**
  Erqiao Lin, Xubu Ma and Chunxiu Qin (Xidian Univ.))
- **Perception of AI Bureaucracy on a Value-Rational and Public Communication Proposals**
  Jinsoon Song (Donga Univ.)
- **Information Sensitivity in the AI-Driven Digital Society: Constructs and Measurement Instrument**
  Xi Chen, Hongying Du (KMMU), Jian Mou (Pusan National Univ.), and Pengxin Zheng (Yunnan Univ.)
- **When Citizens Become Consumers of AI: Algorithm Aversion, Appreciation and Compliance Behavior in AI-Driven Public Health Warning Services**
  Hsiao-Ting Tseng (National Central Univ.)
- **An Exploratory Framework for Flagging Greenwashing: Combining NLP-Based Patent Analysis with Financial-Patent Efficiency Metrics**
  Wen Yan Toh, Wenting Liu((Singapore Univ. of Social Sciences), and Nat Liang (Independent Researcher)

C2 Digital Media, Advertising & Content / Session Chair : Jee-Hae Lim (Univ. of Hawaii at Mānoa) 13:20~14:40

- **Who Tells It Better? Aligning Creation Modalities with Types of PSAs to Enhance Public Attitudes via Systematic Thinking**
  Xiaoxiao Gong and Danfeng Yuan (GuiZhou Univ.)
- **Can Machines Read Your Face? A Video Analytics Framework for Measuring Empathy via Emotional Mimicry in Video Recordings**
  Li Cui (The Hong Kong Polytechnic Univ.), Jee-Hae Lim (Univ. of Hawaii at Manoa), Zhoudao Lu (Univ. of Oxford), Ka Chung Ng, and Jingran Zhao (The Hong Kong Polytechnic Univ.)
- **A Cost-Sensitive Information Grading Model Considering Asymmetric Misclassification Risks**
  Seojin Lee, Yeji Choi, Wonyoung Kim, Hyeseong Lee and Hangbae Chang (Chung-Ang Univ.)
- **How AIGC Disclosure Labels Drive Consumers' UGC Compensatory Search Behavior**
  Zhenhan Qin, Xusen Cheng, Xinhe Qu, Ang Zeng and Yijun Yan (Renmin Univ. of China)
- **Analysis of the Influences of AI on the Media Industry with Story Telling**
  Hui Liu and Jinhwa Kim (Sogang Univ.)
- **Intrusive or Effective? The Trade-Off between Brand Recall and Viewer Satisfaction of In-Game Advertising in eSports**
  Sze Ming Chow, Yan Sun and Sung-Byung Yang (Kyung Hee Univ.)

C3 Digital Finance & Platform Economy / Session Chair : Kyung Young Lee (Dalhousie Univ.) 14:50~16:10

- **Understanding the Endorsement Effect of Digital Virtual Human on Investment Intention Towards Crowdfunding—Based on Elaboration Likelihood Model**
  Jiajie Chu (Southwestern Univ. of Finance and Economics)
- **Probabilistic Earnings Forecasting Using Large Language Models in Digital Platforms**
  Xiaolin Gan, Xincan Qiu and Chenglong Zhang (Kookmin Univ.)
- **Analysis of Solana Validator Staking Amount: A Machine Learning Approach**
  Ke Li and Chenglong Zhang (Kookmin Univ.)
- **Service Fee in C2C Secondhand Markets: Optimal Fee Design for Platform Governance Value**
  Guojun Ji, Chang Liu(Xiamen Univ.), and Kim Hua Tan (Nottingham Univ)
- **How AI Credit Scoring Influences Consumer Satisfaction and Purchase Intention in AI-Driven E-Commerce: Evidence from Bibli PayLater**
  Sheila Silvia Permatasari and Syuan Yi Jiang (National Cheng Kung Univ.)
- **Does the Adoption of Digital Wealth Management Platforms (DWPs) Result in Positive Net Benefits? Retail Investors' Perspective on DWP Success**
  Yong Yin, Kyung Young Lee, Bo Yu, Colin Conrad and Thomas Kb Koo (Dalhousie Univ.)
- **Generative AI-Enabled Value Creation in Robo-Advisory: A Grounded Theory Multiple-Case Study**
  Qin Gao, Weihui Li and Yongqing Yang (Shandong Technology and Business Univ.)

C4 Human Interaction & Future Intelligent Environments / Session Chair : Wenlong Zhu (Qingdao Univ. of Technology) 16:20~17:40

- **How Does Human-AI Fit Shape Consumers' Passive Decision-Making?**
  Wenlong Zhu and Lei Liu (Qingdao Univ. of Technology)
- **Building Trust and Ties: Why Elderly Patients Choose Online Health Communities**
  Jinxin Yang (Hefei Tech. College) and Xinyi Xie (King's College of London)
- **BhashaBridge: An Emotion-Aware Real-Time Indian Sign Language Translator with Offline Doctor Mode Using MediaPipe and LSTM**
  Er.Gautam Kumar, Surbhi Kumari, Roshni Soni, Isha Rani and Manavjeet Singh (Chandigarh Univ.)
- **A Novel Lightweight Medical Image Encryption Scheme Using Convolutional Autoencoders and Chaotic Systems**
  Shaik. Kashif Hussain, Pedaballe Lekha Reddy, Kuruba Kullai Swamy and V. Kowshik Chowdary (Rajeev Gandhi College)
- **Experiencing Joint Agency and Health Self-Efficacy in AI Healthcare: How Perceived Agency Influences Continuance Intention Under Two Dimensions of Anthropomorphism**
  Xiaoan Li and Jing Zhang (Kookmin Univ.)
- **A Story Mining Approach to Solve Problems: A Case on Bridging Metaverse and Elderly**
  Hyunsung Jung, Soyeon Lee and Jinhwa Kim (Sogang Univ.)
- **Research on the Impact of Human-AI Interaction Types on Individual Innovation Performance: A Dual-Pathway Model of Responsibility Attribution and Identification**
  Yujuan Zheng, Shuangwu Dong (XIdian Univ.) and Hanzhuo Ma (Deakin Univ.)
- **Permission Structures, Inferential Processing and Perceived Power Asymmetry in Humanoid Home Robot Contexts**
  Jinjing Li, Jingshi Jin (Kyung Hee Univ.), Kyung Young Lee (Dalhousie Univ.) and Sung-Byung Yang (Kyung Hee Univ.)
- **Embodied AI and Environmeoid Robots Reshape Privacy Calculus in Home Environments**
  Jingshi Jin, Jinjing Li and Sung-Byung Yang (Kyung Hee Univ.)

KITS·ICEC·KIISS
2026 통합 국제학술대회

C1 Responsible AI & Digital Governance

Session Chair : Wenting Liu
(Singapore Univ. of Social Sciences)



C1

Exploring Ethical Challenges of Artificial Intelligence in Commercial Contexts

Ling-Pu Huang, Chen-Hao Huang (National Cheng Kung University)

The rapid advancement and widespread adoption of artificial intelligence (AI) have significantly transformed the ways businesses operate. While the increasing application of AI systems offers substantial efficiency and performance benefits, it also raises complex ethical concerns related to bias, accountability, transparency, and social impact. However, existing studies often address these issues in isolation, without providing an integrative understanding of how such concerns collectively shape organizational practices and digital responsibility. Therefore, this study conducts a systematic review to examine ethical issues associated with AI adoption in business contexts. The research adopts a two-stage analytical design combining Main Path Analysis (MPA) and Latent Dirichlet Allocation (LDA). This paper presents the findings from the MPA stage, which reveal a three-phase evolutionary progression and two parallel development paths. These findings help clarify current ethical issues related to the development and deployment of AI systems in commercial contexts.

KEY WORDS

AI ethics, Artificial Intelligence, Ethic, GenAI, Systematic Review

C1

How AI Transparency Affects AI Appreciation: The Mediating Roles of Trusting Beliefs

Hua Liu (School of Business, Anhui University),
Xiaofan Tang (School of Health Management, Anhui Medical University),
Jinmei Yin (School of Economics and Management, Nanjing University of Aeronautics and Astronautics)

This paper examines how AI transparency influences students' AI appreciation through the mediating role of trusting beliefs, drawing on the knowledge-based trust perspective. This study proposes that trusting beliefs-comprising competence, integrity, and benevolence-mediate the relationship between AI transparency and AI appreciation. Using the "judge-advisor system" paradigm, we conduct experiment designed within the Xuexitong app context to test these hypotheses. This research contributes to understanding how transparency builds knowledge-based trust and enhances human-AI collaboration in educational environments.

KEY WORDS

AI transparency, trusting beliefs, AI appreciation

C1

AI-Driven Intelligent Policy Monitoring Method: Measuring the Intensity and Synergy of China's Data Element Market

Erqiao Lin, Xubu Ma, Chunxiu Qin (Xidian University)

As the global economy enters the AI Transformation (AX) era, data elements have become the fundamental energy driving AI evolution. However, the governance of data markets often faces a "black box effect," where a significant "knowledge-action misalignment" exists between ambitious policy goals and limited execution tools. This study proposes PISA (Policy Intelligence & Synergy-oriented Architecture), an agent-like framework that simulates the "Perception-Cognition-Modeling-Feedback" loop to achieve automated, structural monitoring of policy texts.By analyzing 2,911 policy articles from China (2021~2023), we find that policy synergy, rather than single-dimensional intensity, is the core engine for market growth. While most regions exhibit "Synergy Gain," mature AX ecosystems show "Market Spontaneity" where market evolution outpaces policy iteration. This research provides a scalable, AI-driven evaluation tool for deciphering the transmission mechanism from institutional intent to market output, offering a new digital governance paradigm for the global AX ecosystem

KEY WORDS

AI Transformation (AX), Policy Intelligence, Agent-like Framework

C1

Perception of AI Bureaucracy on a Value-Rational and Public Communication Proposals

Jinsoon Song (Dong-a univ. Public ADMINISTRATION)

This study examines pre-service public officials' perceptions (N=169) of AI Bureaucracy, focusing on how instrumental rationality and mechanical impersonality affect expectations for democratic participation. Descriptive analysis reveals a significant concern regarding the black-box nature of AI, with 87.5%of respondents fearing decreased civic participation (Mean=4.42). Multiple regression analysis yielded two pivotal findings: 1) Mechanical Impersonalityis the strongest predictor of expectations for Public Communication (β =.512, p<.001), suggesting that AI's perceived lack of empathy triggers a normative demand for reinforced democratic engagement. 2) Instrumental Consistency significantly drives the need for Ethical Accountability (β =.428, p<.001), highlighting the demand for human-in-the-loop oversight. These results emphasize that AI governance must transcend technical efficiency by prioritizing algorithmic explainability and human-mediated communication to ensure administrative legitimacy in the digital era.

KEY WORDS

AI Bureaucracy, Mechanical Impersonality, Public Communication

Exploring Pre-service Public Officials' Perceptions of AI Bureaucracy: A Value-Rational Perspective and Public Communication Proposals

Song, J.S
Department of Pubic Administration, Donga University
E-mail: songjs@dau.ac.kr

Abstract

This study examines pre-service public officials' perceptions (N=169) of AI Bureaucracy in South Korea, focusing on how instrumental rationality and mechanical impersonality affect expectations for democratic participation. Descriptive analysis reveals a significant concern regarding the black-box nature of AI, with 87.5% of respondents fearing decreased civic participation (Mean=4.42). Multiple regression analysis yielded two pivotal findings: 1) Mechanical Impersonality is the strongest predictor of expectations for Public Communication (β=.512, ρ<.001), suggesting that AI's perceived lack of empathy triggers a normative demand for reinforced democratic engagement. 2) Instrumental Consistency significantly drives the need for Ethical Accountability (β=.428, ρ<.001), highlighting the demand for human-in-the-loop oversight. These results emphasize that AI governance must transcend technical efficiency by prioritizing algorithmic explainability and human-mediated communication to ensure administrative legitimacy in the digital era.

Keywords: AI Bureaucracy, Mechanical Impersonality, Public Communication, Democratic

Introduction

The field of public administration is undergoing a fundamental transformation due to the integration of AI, a core technology of the Fourth Industrial Revolution. AI is spearheading innovations in administrative processes by analyzing vast amounts of public data, supporting complex decision-making, and automating public services. This burgeoning role of AI raises novel questions regarding the essence of bureaucracy, a long-standing discourse in public administration. Specifically, the introduction of AI is maximizing instrumental rationality (Zweckrationalität), one of the core values of bureaucracy proposed by Max Weber(Weber, 1978). By providing consistent and predictable outcomes devoid of human emotion or discretion, AI's algorithm-based processing implements rule-based rationality in the digital era, ushering in the age of algorithmic bureaucracy. However, when AI remains confined to mere efficiency—namely, instrumental rationality—conflicts with value rationality (Wertrationalität), the other pillar of public administration, become inevitable[1]. The "black-box problem" of AI systems hinders transparency in decision-making, and inherent biases in training data can undermine the fairness of services. These issues pose significant threats to accountability and ethics, the foundations of administration. Consequently, modern public administration faces the fundamental challenge of integrating AI not merely as a tool for efficiency, but as a new form of bureaucracy that realizes public values. The purpose of this study is to empirically investigate the perceptions of South Korean public administration students regarding AI's bureaucratic characteristics based on Weber's dual structure of rationality and to identify how these perceptions affect value-rational expectations toward AI. To this end, this study performs two core tasks: First, it analyzes how pre-service officials perceive AI's rule-based consistency and impersonality in terms of maximizing instrumental rationality. Second, it empirically identifies the influence of these perceptions on value-rational expectations, such

as ethical responsibility and transparency. The significance of this research lies in extending Max Weber's bureaucratic theory to the AI era and proposing a new analytical framework that explores AI as a novel tool for value-rational bureaucracy, rather than a simple instrument. This will contribute to exploring the feasibility of administrative values without viewing AI's impact through a simplistic binary lens. Furthermore, based on the perceptions of future public officials, it emphasizes that securing ethical responsibility and transparency in AI design and operation is the path to enhancing the value-rationality of public administration. This paper is organized into five chapters. Chapter 2 reviewed the theoretical background, focusing on the dual structure of Weberian rationality, the bureaucratic characteristics of AI, and the concept of value-rational expectations, and establishes research hypotheses. Chapter 3 described the research measurement tools and survey results. Chapter 4 conducted an empirical analysis based on the survey results and offers suggestions for public communication. Finally, Chapter 5 concluded this paper's findings and directions.

2. Theoretical Background

2.1. Max Weber's Bureaucracy and the Dual Structure of Rationality

When asking a layperson about AI, names like Apple's Siri, Amazon's cloud services, Tesla's autonomous vehicles, or Google's search algorithms are often invoked. If AI systems are deemed more reliable or rational than human experts, it suggests that they should be entrusted with critical decisions in health, education, and criminal justice. However, focusing solely on specific algorithmic techniques often emphasizes continuous technological advancement while neglecting the broad environmental impacts and computational costs on a strained planet. In *Economy and Society*, Max Weber proposed that bureaucracy is the most rational form of government, establishing its essential role in ensuring legal order through administration. Weber's logic centers on control and legal certainty, exemplified by how the relationship between the administration and the governed is mediated through common administrative processes. Ultimately, Weber's ideal-typical bureaucracy operates independently of the dynamics of politics and politicians; within this structure, authority is decentralized, delegated, clearly delineated, and hierarchical(Weber, 1978). Weber(1978) distinguished between two core types of rationality to understand social action and domination, which serve as the theoretical foundation for this study's analysis of AI administration. First, instrumental rationality (*Zweckrationalität, IR*) refers to actions where means and ends are rationally calculated and selected to achieve a specific goal. In a bureaucracy, this manifests as rule-based processing that derives the most predictable results in the most efficient manner. The introduction of AI is evaluated as maximizing this instrumental rationality through its rapid data processing and capacity for error-free, repetitive execution. AI acts as a tool that achieves set objectives (e.g., welfare recipient selection, traffic control) via optimal methods, excluding the subjective discretion or emotions of human bureaucrats. Second, value rationality (*Wertrationalität, VR*) involves actions pursued for their intrinsic value—such as ethics, justice, honor, or duty—regardless of the outcome. In the administrative sphere, this means treating public values like fairness, transparency, accountability, and democracy as ends in themselves. Weber warned that the pursuit of instrumental rationality could transform modern bureaucracy into an Iron Cage (*Stahlhartes Gehäuse*), where human values and meaning are lost(Weber, 1946). This caution aligns with the modern risk of technological efficiency addiction inherent in AI. This study moves beyond AI's role in enhancing instrumental rationality to analyze how the way pre-service public officials perceive AI's bureaucratic characteristics stimulates value-rational expectations (demands for ethics and transparency), exploring the potential for AI to become a partner in realizing values rather than a mere tool.

2.2. Algorithmic Bureaucracy vs Value-Rational Expectations

AI bureaucracy refers to a system that automates and streamlines public service delivery and administrative decision-making based on data and algorithms. It offers the advantage of resolving traditional bureaucratic inefficiencies and enabling swift, objective decisions. As AI integrates into public administration, it gives rise to the concept of Algorithmic Bureaucracy, which applies Weberian concepts to the digital environment(Bovens and Zouridis, 2002). In an algorithmic bureaucracy, administrative functions are performed by rule-based algorithms, implementing the core characteristics of Weberian bureaucracy through digital automation and standardization. IR most closely linked to AI are instrumental consistency and mechanical impersonality. This study explores how perceptions of

these two characteristics among public administration students influence their value-rational expectations toward AI. VR expectations for current and future AI administration encompass public values that must accompany the pursuit of efficiency, namely, administrative ethics and public communication. These are measured through Perception of Ethical Responsibility and Demand for Transparency and Public Communication. From an administrative perspective, for AI-mediated services to gain public trust and legitimacy, they must move beyond one-way information provision toward establishing two-way communication channels, such as grievance procedures for AI decisions. Currently, self-regulatory ethical frameworks allow corporations to choose how they deploy technology and define ethical AI(Song, 2023). Tech companies rarely face serious financial penalties or sanctions when AI systems violate laws or ethical principles. Furthermore, publicly traded companies often prioritize maximizing investor returns over ethical concerns, rendering ethics secondary to profit. Finally, as AI bureaucracy increasingly automates administrative discretion, it is essential to recognize that the current era demands a resurgence of humanistic and philosophical reflection to ensure the legitimacy of public values.

3. Research Methodology and Analysis of Results

3.1. Research Design

This study employs a quantitative research design to empirically investigate how pre-service public officials perceive the bureaucratic nature of AI and how these perceptions influence their value-rational expectations. The sample consists of 169 undergraduate students majoring in Public Administration who possess a foundational understanding of bureaucratic theory and administrative values. Data were collected through a structured survey to capture the distribution of perceptions regarding the structural and ethical impacts of AI implementation. To ensure the reliability of the constructs, Cronbach's α analysis was conducted, with all variables exceeding the 0.70 threshold (range: 0.703 to 0.824), confirming high internal consistency. The data were analyzed using Descriptive Statistics to identify general perception trends, Correlation Analysis to examine the relationships between variables, and Multiple Regression Analysis to determine the relative influence of AI bureaucratic perceptions on value-rational expectations(Table 7).

Based on Max Weber's theory of bureaucracy, the survey instrument was structured into three core dimensions:

1. Instrumental Consistency: Defined as the degree to which AI follows a predictable and rigid logic similar to legal regulations, ensuring error-free execution of repetitive tasks. This was measured by assessing respondents' perceptions of AI's adherence to programmed algorithms and its impact on administrative flexibility. Measures the dilemma between AI's predictability/efficiency and the potential loss of flexibility (items 2-1 to 2-4).
2. Mechanical Impersonality and Objectivity (Bias): Defined as the human-absent nature of algorithmic processing, characterized by a lack of empathy and contextual understanding. This dimension measures the psychological distance and coldness perceived in AI-driven administrative dispositions. Evaluates perceptions of whether AI's impersonal processing solidifies data bias or leads to dehumanized administrative experiences (items 3-1 to 3-4).
3. Ethical Accountability and Control: Assesses perceptions of shifts in control and responsibility, including the ambiguity of accountability during AI errors, the risk of administrative centralization, and the reduction of civil servants' discretionary power (items 4-1 to 4-4).
4. Public Communication & Transparency: This represents the demand for communicative rationality ensuring that AI decision-making is not a "black box" but is explainable and open to civic participation. It measures the requirement for agencies to disclose algorithmic logic to maintain democratic legitimacy(items 5-1 to 5-4/item 6-1 to 6-4).

The analytical procedure followed a stepwise approach to ensure causal coherence:

- First, Descriptive Statistics were used to identify the general characteristics of respondents and the basic distribution of their perceptions regarding the utility and risks of AI bureaucracy.
- Second, Correlation Analysis was conducted to examine the direction and strength of the relationship between the independent variables (AI bureaucratic characteristics) and dependent variables (value-rational expectations), as well as to check for multicollinearity (see Table7).
- Third, Multiple Regression Analysis was performed as the core analytical stage to identify the relative influence of perceptions of AI characteristics on various sub-factors of value-rational expectations (see Table 8). Statistical significance was determined based on a p-value of < 0.05.

3.2. Analysis of Results

Table 1. Demographic Characteristics and AI Usage Status of Respondents (N=169)

Category	Sub-category	Frequency(n)	Percentage(%)
Grade	Sophomore(2^{ND} Year)	75	44.4
	Junior(3^{RD} Year)	54	32.0
	Senior(4^{TH} Year)	20	23.6
Interest in AI	Very High	35	20.8
	High	42	25.0
	Moderate	77	45.8
	Low / Very Low	15	8.4
Purpose of AI Use*	Data Collection & Analysis	80	47.4
	Academic Study & Writing	62	36.8
	Everyday Information Search	27	15.8
Frequency of AI Use	Daily	11	6.3
	2–3 times per week	49	29.1
	Once per week	60	35.4
	Less than once a month / Never	49	29.1

As presented in Table 1, the survey participants consisted of 169 undergraduate students majoring in Public Administration. A majority of the respondents were sophomores (54.2%) and juniors (25.0%), indicating that the sample represents a pre-expert group with foundational knowledge of administrative theories and bureaucratic systems. Regarding interest in AI, 45.8% of students reported a high level of interest, while more than 70% of the total respondents engaged with AI tools at least once a week. These empirical experiences provide a cognitive basis for respondents to evaluate the technical attributes of AI in relation to administrative principles.

3.2.2. Multidimensional Analysis of Perceptions Toward AI Bureaucratic Characteristics.

This section analyzes how Weberian bureaucratic characteristics manifest through algorithms, focusing on three core elements: rule-based consistency and rigidity, impersonality and bias, and accountability and control structures.

Table 2. Perceptions of Instrumental Consistency and Rigidity

Item No.	Survey Statement	Disagree / Strongly Disagree (%)	Neutral (%)	Agree / Strongly Agree (%)	Mean (SD)*
2-1	AI data processing follows a predictable and consistent logic, similar to legal regulations.	29.2	45.8	25.0	2.92(0.87)
2-2	AI responds mechanically by strictly adhering to programmed algorithms without modification.	22.9	39.6	37.5	3.19(0.91)
2-3	Decisions made by AI will be more objective as they exclude human emotions or personal feelings.	14.6	31.2	54.2	3.52(1.05)
2-4	While AI implementation increases processing speed, it will reduce creative and flexible solutions.	10.4	27.1	62.5	3.79(0.98)

Regarding Table 2, the survey examined whether AI data processing follows predictable and consistent logic similar to legal regulations (item 2-1). The results showed that 45.8% of respondents answered "Neutral," while only 25.0% provided positive responses (4 or 5 on the Likert scale). This indicates a prevailing neutral stance rather than a firm conviction regarding AI's consistency. Similarly, for the item stating that AI responds mechanically by strictly adhering to algorithms without modification (item 2-2), 39.6% of responses were concentrated in the Neutral category, suggesting that perceptions of AI's rigidity are not yet definitively formed. However, a significant majority of respondents (62.5%) agreed with the statement that While AI implementation will increase processing speed, it will reduce creative and flexible solutions (item 2-4). This confirms a dominant perception that AI will simultaneously enhance bureaucratic efficiency (pro) while exacerbating administrative rigidity (con). In essence, while pre-service officials expect AI to adhere strictly to rules in a manner consistent with traditional bureaucracy, they express high levels of concern that such adherence will inevitably lead to a loss of administrative flexibility (table 3-2).

Table 3. Perceptions of Mechanical Impersonality and Bias

Item No.	Survey Statement	Disagree / Strongly Disagree (%)	Neutral (%)	Agree / Strongly Agree (%)	Mean (SD)*
3-1	AI processing of tasks (e.g., civil petitions, welfare eligibility) lacks human empathy and contextual understanding.	8.3	20.8	70.9	3.88(0.84)
3-2	Algorithmic bias in AI learning data is likely to be fixed in public services in the form of objective rules.	12.5	33.3	54.2	3.58(0.92)
3-3	AI-handled tasks should be limited to simple duties that do not require human empathy or contextual insight.	16.7	16.6	66.7	3.81(1.02)

3-4	Administrative dispositions by AI (e.g., license revocation, fines) will feel colder and more impersonal than those by human officials.	6.3	25.0	68.7	3.94(0.89)

The findings in Table 3 reveal a high level of concern regarding AI's mechanical impersonality. Notably, 70.9% of respondents agreed that AI lacks human empathy and contextual understanding (Item 3-1), and 68.7% felt that AI-driven administrative decisions would be perceived as 'colder' than those made by humans (Item 3-4). This reflects the 'iron cage' of algorithmic bureaucracy (Weber,1978), where technical objectivity may result in human alienation. Furthermore, a majority (66.7%) emphasized that AI should be restricted to simple tasks (item 3-3), suggesting that pre-service officials prioritize value-rational judgment over instrumental efficiency in complex administrative contexts.

Table 4. Perceptions of Accountability and Control Structures

Item No.	**Survey Statement**	**Disagree / Strongly Disagree (%)**	**Neutral (%)**	**Agree / Strongly Agree (%)**	**Mean (SD)***
4-1	In the event of an AI error causing harm to citizens, it is difficult to determine whether responsibility lies with the developer or the approving official.	14.6	27.1	58.3	3.65(0.93)
4-2	Procedures for appealing AI decisions will be more complex and inefficient than current administrative litigation processes.	25.0	41.7	33.3	3.33(0.88)
4-3	The integration of AI systems is likely to intensify administrative centralization of local governments by the central government.	10.4	45.8	43.8	3.50(0.87)
4-4	Future civil servants will experience a reduction in their professional discretion as they become more dependent on AI-driven decisions.	12.5	29.2	58.3	3.63(0.96)

As illustrated in Table 4, respondents expressed significant concerns regarding the erosion of accountability. Notably, 58.3% of pre-service officials agreed that identifying the locus of responsibility for AI errors is challenging (item 4-1), highlighting the emergence of the 'problem of many hands' in algorithmic governance. Furthermore, another 58.3% perceived that AI would lead to a reduction in administrative discretion (item 4-4). This suggests that while AI is implemented to enhance efficiency, it may paradoxically weaken the ethical agency of individual bureaucrats. The perception that AI promotes centralization (47.9%, item 4-3) further indicates that future administrators view AI not just as a neutral tool, but as a mechanism that could reshape the power dynamics within the bureaucratic hierarchy.

Table 5. Expectations for Public Communication and Democratic Participation

Item No.	**Survey Statement**	**Disagree / Strongly Disagree (%)**	**Neutral (%)**	**Agree / Strongly Agree (%)**	**Mean (SD)***

5-1	AI systems could hinder multi-layered communication by providing uniform and standardized information.	4.2	16.7	79.1	4.02(0.76)
5-2	Disclosure of AI algorithmic logic and data processing is essential for administrative democratic legitimacy.	6.3	25.0	68.7	3.85(0.85)
5-3	AI intervention would decrease opportunities for civic participation in policy-making due to "black-boxing."	2.1	10.4	87.5	4.42(0.68)
5-4	Expansion of AI-driven administration could exacerbate the service gap for digitally vulnerable populations.	4.2	18.8	77.0	4.08(0.79)

The findings presented in Table 5 underscore a profound commitment to the democratic values of public administration among pre-service officials. Most notably, 87.5% of respondents, the highest agreement rate in this category, insisted on the reinforcement of offline channels for digitally vulnerable groups (item 6-3). This reflects a critical awareness of the 'digital divide' and a demand for inclusive governance. Furthermore, 79.1% emphasized the necessity of human-mediated consultation (item 6-1), and 77.0% agreed that AI's ultimate purpose should be the promotion of democratic citizenship rather than mere administrative convenience (item 6-4). These results suggest that future administrators do not view AI as a replacement for democratic discourse; rather, they demand that AI be subservient to the overarching values of social equity and communicative rationality.

Table 6. Expectations for Administrative Ethics and Transparency

Item No.	**Survey Statement**	**Disagree / Strongly Disagree (%)**	**Neutral (%)**	**Agree / Strongly Agree (%)**	**Mean (SD)***
6-1	Even if AI handles the logic, the final ethical and legal responsibility for administrative results must rest with a "human" official.	10.4	31.3	58.3	3.65(0.93)
6-2	Public agencies must provide citizens with clear explanations regarding the operational logic and data sources of AI algorithms.	14.6	41.7	43.7	3.33(0.88)
6-3	To prevent algorithmic bias, an independent external organization should regularly audit the fairness of AI administrative systems.	8.3	43.8	47.9	3.50(0.87)
6-4	Enhancing the "Explainability" of AI is more important than simply increasing the speed or efficiency of administrative processing.	12.5	29.2	58.3	3.63(0.96)

The data in Table 6 highlights a robust demand for value-rational safeguards. Notably, 85.4% of respondents emphasized the necessity of disclosing algorithmic logic (item 6-2), and 81.2% insisted that final accountability must remain with human officials (item 6-1). This suggests that as AI takes over instrumental tasks, pre-service officials increasingly prioritize 'Explainability' and 'Human-in-the-loop' principles over mere technical efficiency. These high agreement rates reflect a collective

recognition that the legitimacy of AI-driven administration is contingent upon its ethical transparency rather than its operational speed.

4. Empirical Analysis: Testing the Impact of AI Bureaucratic Perceptions

4.1. Reliability and Correlation Analysis

To ensure the internal consistency of the constructs, Cronbach's α was calculated for each variable. All variables exceeded the recommended threshold of 0.70, demonstrating high reliability: Mechanical Impersonality (0.824), Instrumental Consistency (0.765), Ethical Accountability (0.748), and Public Communication (0.703). The correlation matrix reveals significant positive relationships among the variables. Notably, a strong correlation exists between Mechanical Impersonality and Ethical Accountability ($\gamma = 0.512$, $\rho < 0.01$). This suggests that as future officials perceive AI as more mechanical and devoid of empathy, their demand for rigorous ethical responsibility increases. Additionally, Public Communication shows a significant positive correlation with all other perceptions, indicating that these bureaucratic views are intrinsically linked to expectations for democratic engagement.

Table 7. Correlation Matrix and Reliability Results

Variables	1	2	3	4	Mean	SD	α
Instrumental Consistency	1				3.10	0.91	.765
Mechanical Impersonality	.342*	1			3.80	0.86	.824
Ethical Accountability	.428**	.512**	1		4.15	0.76	.748
Public Communication	.289*	.458**	.394*	1	4.22	0.72	.703

*ρ < .05, **ρ<.01

To verify the internal consistency of the measurement scales, Cronbach's α coefficients were calculated. All variables exceeded the threshold of 0.70, indicating high reliability for further analysis. Specifically, Mechanical Impersonality showed the highest reliability at 0.824, followed by Instrumental Consistency (0.765) and Ethical Accountability (0.748). Even the outcome variable, Public Communication, demonstrated acceptable consistency at 0.703. The mean scores provide insight into the normative priorities of pre-service officials (N=169). The highest mean was observed in Public Communication (4.22), followed by Ethical Accountability (4.15). This confirms that respondents place the highest value-rational importance on democratic engagement and ethical responsibility. In contrast, Instrumental Consistency recorded the lowest mean (3.10), suggesting a relatively neutral or skeptical view toward the technical consistency of AI compared to its ethical implications. The correlation matrix reveals significant positive relationships among all variables, supporting the study's theoretical framework:

- Impersonality and Ethics: The strongest correlation was found between Mechanical Impersonality and Ethical Accountability ($\gamma = 0.512$, $\rho < 0.01$). This suggests that as future officials perceive AI as more mechanical and cold, their demand for rigorous ethical safeguards increases proportionally.
- Consistency and Ethics: Instrumental Consistency also showed a moderate correlation with Ethical Accountability ($\gamma = 0.428$, $\rho< 0.01$), indicating that the more AI is seen as a rule-following logic engine, the more respondents emphasize the need for human-centric accountability.
- Impact on Public Communication: Public Communication was significantly correlated with Mechanical Impersonality ($\gamma= 0.458$, $\rho< 0.01$). This is a key finding, implying that those who

are most concerned about AI's lack of empathy are also those who most strongly advocate for reinforced democratic participation and communication channels.

4.2. Multiple Regression Analysis: Driving Value-Rational Expectations

To empirically verify the causal pathways, multiple regression analyses were performed. The results are summarized in Table 8.

Table 8. Multiple Regression Results on Value-Rational Expectations

Independent Variables	**Model 1: Ethical Accountability(β)**	**Model 2: Public Communication(β)**
Instrumental Consistency	.428***	-.085
Mechanical Impersonality	.154*	.512***
Accountability& Control	.295**	.142
R-squared(R^2)	.485	.392
F-value	15.42***	11.24***

* ρ< .05, ** ρ< .01, *** ρ< .001

As shown in Table 8, the multiple regression analysis reveals how perceptions of AI's bureaucratic characteristics influence two distinct dimensions of value-rational expectations among pre-service public officials.

1. Predictors of Ethical Accountability (Model 1)

In Model 1, Instrumental Consistency (β =0.428, ρ <0 .001) emerged as the most significant predictor. This indicates that as respondents perceive AI to be more rule-governed and predictable, their demand for ethical accountability and transparency increases likely as a safeguard against algorithmic rigidity. Additionally, Structural Controllability (β = 0.295, ρ< 0.01) and Mechanical Impersonality (β= .154, ρ < 0.05) also showed significant positive effects. The model explains 48.5% of the variance (R^2 = 0.485, F = 15.42).

2. Predictors of Public Communication and Participation (Model 2)

In Model 2, Mechanical Impersonality (β= 0.512, ρ< 0.001) was the only dominant and significant predictor. This suggests that the "cold" and dehumanized nature of AI-driven administration is the primary driver behind the demand for empathetic, human-mediated communication and democratic participation. Interestingly, Instrumental Consistency and Structural Controllability did not significantly influence communication demands. This model accounts for 39.2% of the variance (R^2 = 0.392, F = 11.24, ρ< 0.001). In contrast, Instrumental Consistency (β= -0.085) did not show a statistically significant positive impact in this model.

4.3. Perceptions of Communication and Participation

As shown in Table 3-5, regarding the impact on communication, nearly half of the respondents (79.1%) agreed that AI systems could hinder multi-layered communication by providing information in a uniform and standardized manner. While automated response systems maximize rule-based consistency for speed, they represent an extreme manifestation of bureaucratic impersonality. Respondents voiced concerns that the lack of contextual understanding in AI could lead to an emotional and qualitative disconnection between the government and its citizens. Furthermore, there was an overwhelming demand for transparency(68.7%), with respondents asserting that disclosing AI

algorithmic logic is essential for democratic legitimacy. Transparency is viewed as the minimum safeguard to ensure fairness and accountability. This underscores the policy necessity of mandating Explainable AI (XAI) principles, ensuring that data sources and processing logic are disclosed to the public. Regarding civic participation, 87.5% of respondents feared that AI intervention would reduce opportunities for policy involvement, as decision-making processes become "black-boxed." Lastly, the findings highlight concerns over the digital divide(77.0%), where AI-driven services might act as a barrier for vulnerable groups (the elderly, low-income, etc.). This implies that AI's mechanical impersonality could transcend mere "coldness" and evolve into systemic discrimination, ultimately deepening administrative inequality.

4.4. Discussion-Governance Recommendations for Mitigating AI Bureaucracy

The empirical results reveal a Value-Rational response among future public officials. While AI's Instrumental Consistency drives the need for Ethical Accountability, it is the Mechanical Impersonality of AI that triggers a powerful demand for Public Communication and Democratic Participation. This suggests that to successfully integrate AI into the public sector, policymakers must address the human deficit of algorithms. The findings advocate for a Hybrid Governance Model where AI's efficiency is checked by human-centric communication channels and transparent participation mechanisms to maintain administrative legitimacy. To transcend the limitations of algorithmic rigidity, this paper proposes a Hybrid Governance Model that harmonizes technical optimization with democratic accountability. In this model, AI serves as an Augmentative Intelligence that processes complex data structures to enhance administrative speed. Concurrently, human bureaucrats act as Value Guardians, exercising discretionary judgment to ensure that automated outputs align with constitutional values and social equity. The core of this hybridity lies in the Strategic Decoupling of Tasks while the AI handles the *instrumental logic* of how to achieve a goal, the human oversight mechanism focuses on the *value-rational logic* of why a certain action is being taken. This symbiotic relationship creates a Check and Balance system within the digital bureaucracy, preventing algorithmic alienation and ensuring that technology remains a tool for public service rather than a master of public policy.

5. Policy Recommendations for Public Communicative

5.1. The Necessity of Value-Rational Oversight in the AI Era

The pervasive integration of AI in daily life may enhance operational efficiency, but algorithmic opacity remains a critical threat that can amplify social harms and bias. As illustrated by the "Filter Bubble" phenomenon, algorithms often confine users within ideological silos, exacerbating polarization. High-profile whistleblowing cases, such as those involving Facebook, have revealed how secretive algorithmic practices can jeopardize child safety and incite racial violence(C-SPAN 2021). Furthermore, extreme phenomena like AI psychosis—exemplified by the tragic suicide of a minor obsessed with an AI chatbot—demonstrate that technology companies cannot be left to self-regulate(Hankyoreh Daily 2024). Therefore, public administration must view AI not merely as a tool but as a digital bureaucrat requiring rigorous ethical and institutional control.

This study explored the perceptions of 169 pre-service public officials regarding the transition toward AI-driven bureaucracy. The empirical analysis yielded several critical insights:

- Prevalence of Value-Rational Concerns: Future officials are more concerned with the Mechanical Impersonality and Black-box nature of AI than its technical efficiency.
- Determinants of Accountability: The demand for Ethical Accountability is primarily driven by the perceived Instrumental Consistency ($\beta = 0.428$, $\rho < 0.001$) of AI, suggesting that as systems become more rigid, the need for human ethical intervention becomes more acute.
- Drivers of Democratic Participation: The most significant predictor for Public Communication and Participation is Mechanical Impersonality ($\beta = 0.512$, $\rho < 0.001$). This confirms that the emotional and contextual coldness of AI triggers a strong normative demand for reinforced democratic safeguards.

5.2. Strategic Recommendations for AI PA on Public Communication

Based on the empirical findings of this study, the following policy recommendations are proposed:

1. Institutionalizing human-in-the-loop for Ethical Accountability

Given that 58.3% (n=99)(table 6, item 6-1) of respondents insisted that final legal and ethical responsibility must rest with a human official, the government should:

- Establish a clear legal framework that defines the boundaries of AI autonomy and preserves the final decision-making power of human bureaucrats(Song, 2024).
- Implement Internal AI Ethics Committees within public agencies to oversee algorithmic logic and ensure it aligns with constitutional values(Song, 2022).

2. Establishing a Hybrid Communication Model to Bridge the Digital Divide(Song, 2023, 2022).

Based on the overwhelming demand for offline and face-to-face communication (87.5%, n=148)(table 5 item 5-3), policy efforts should focus on:

- Maintaining human-mediated service channels alongside AI systems to protect the rights of digitally vulnerable populations (e.g., the elderly and disabled), (table 5, item 5-4).
- Developing "Omni-channel" administrative services where AI handles routine tasks while human officials handle complex cases requiring empathy and contextual judgment.

3. Enhancing Algorithmic Transparency and Explainability

The empirical link between the "Black-box" effect and decreased civic participation (table 5, item 5-3, Mean=4.42) necessitates:

- The mandatory disclosure of AI logic and data sources for administrative decisions that directly impact citizens' rights.
- The creation of an Independent AI Audit Organization to regularly verify the fairness and neutrality of algorithms, as supported by 47.9% of the respondents (table 3-6, item 6-3).

4. Redesigning Public Official Training Programs

As future officials anticipate a reduction in discretion (58.3%, n=99), education for the next generation of bureaucrats must shift(table 6, item 6-4):

- From simple technical literacy to Digital Citizenship and Value-Rational Judgment.
- Training should emphasize how to interpret AI outputs critically and integrate them with public interest values rather than following algorithms blindly.

The future of AI bureaucracy should not be a choice between efficiency and democracy. As demonstrated by the high expectations for Public Communication (Mean=4.22), the successful integration of AI requires a Value-Centered Governance approach. By addressing the Mechanical Impersonality of AI through reinforced democratic participation and ethical accountability, the public sector can harness the power of technology without losing its essential human character.

6. Conclusion

AI administration must transcend mere instrumental rationality focused on efficiency and must realize public values through value-rational reflection. Ethical reflection on AI is a critical act of reassessing how technological deployment affects fundamental values such as human dignity, social justice, and democracy. Just as Max Weber warned that a bureaucracy pursuing only instrumental rationality could become an Iron Cage devoid of humanity, AI risks devolving into an AI in an Iron Cage depriving humans of discretion and flexibility while perpetuating algorithmic bias if it unthinkingly prioritizes efficiency alone. This study empirically demonstrated that the perceptions of AI’s bureaucratic traits (consistency and impersonality) held by Korean public administration undergraduates significantly influence their expectations for ethical responsibility, transparency, and public communication. The findings suggest that the moment AI maximizes rule-based consistency and impersonality, it inevitably creates risks of rigidity and responsibility avoidance that must be explicitly addressed and needed human-centered control, governance roadmap and strengthening public values

education. In conclusion, humanistic reflection on AI serves as both the starting point and the compass for ensuring that AI administration does not fall into the trap of instrumental rationality. The challenge of an AI bureaucracy is not merely an issue of administrative efficiency, but a fundamental question regarding the nature of the society intended for human habitation. By designing AI systems as tools to realize human values rather than 'cruel optimization,' the discourse can move toward a sustainable model of collective politics that transcends market logic. The essential mission remains to redraw the course from this point forward, prioritizing life over optimization and solidarity over value extraction.

REFERENCE

Bovens Mark and Stavros Zouridis. 2002. "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control". *Public Administration Review* 62(2): 174-184

https://www.jstor.org/stable/3109901

C-SPAN. 2021. "Facebook whistleblower testifies online about child protection." October 5. (Search date:2025.11.15).https://www.c-span.org/program/senate-committee/facebook-whistleblower-testifies-on-protecting-children-online/603426

HankyorehDaily. 2024. "AI Chatbot Leads 14-Year-Old Son to Suicide Litigation With Production Company." October 24 (Search Date: 2025.11.25).

https://www.hani.co.kr/arti/international/international_general/1164119.html

Song, Jinsoon. 2024. "AI Act, 2024 Presidential Election, and Human-AI Public Communication Verification Project to Fight Disinformation." *J Intell Inform Syst 30(3):* 137-159.

DOI: 10.13088/jiis.2024.30.3.137

Song, Jinsoon. 2023. "Ethical Demands in Law Enforcement based on Making Decision of RobotPolice: Black Boxes, Biases and Ethical Governance Framework", *NARS 15(2)*: 80-115. DOI :10.22809/nars.2023.15.2.004

Song, Jinsoon. 2022. "The Introducing voice -based public services for strengthening the accessibility of the social vulnerables and open public communication." *J Intell Inform Syst 28(2)*: 279~306. https://doi.org/10.13088/jiis.2022.28.2.279

Weber, Max. 1978. *Economy and Society*: *An Outline of Interpretive Sociology*. Edited by Guenther Roth and Claus Wittich. University of California Press.

Weber, Max. 1946. *From Max Weber: Essays in Sociology*. H. H. Gerth & C. W. Mills, Eds. & Trans. Oxford University Press.

C1

Information Sensitivity in the AI-Driven Digital Society:Constructs and Measurement Instrument

Xi Chen (YNU), Hongying Du (KMMU), Jian Mou (Pusan National University) and Pengxin Zheng (YNU).

The operation of the AI-driven digital society increasingly relies on the processing of data by platform algorithms. Information sensitivity has become a concept of growing concern in the fields of information security and privacy protection. Existing research remains fragmented regarding the conceptual definition, dimensional delineation, and measurement of information sensitivity, and a systematic and reliable measurement instrument has yet to be established. This study adopts a mixed-methods approach to develop and validate a scale for information sensitivity. It was found that information sensitivity comprises three core dimensions: risk and loss, collection concerns, and user relatedness. The constructed scale demonstrates good reliability and validity, thereby providing an effective measurement tool for subsequent research in this field.

KEY WORDS

AI-driven digital society, information sensitivity, scale development, mixed-methods research

C1

When Citizens Become Consumers of AI: Algorithm Aversion, Appreciation, and Compliance Behavior in AI-Driven Public Health Warning Services

Hsiao-Ting Tseng (National Central University).

AI-driven advisory systems are transforming both e-commerce and public governance, yet when and why citizens follow or override algorithmic recommendations remains poorly understood. Drawing on Algorithm Aversion and Appreciation theory integrated with Protection Motivation Theory (PMT), this research-in-progress examines citizens' behavioral responses to AI-generated air quality warnings. We propose that algorithm attitude, shaped by perceived AI error visibility, warning accuracy history, and AI literacy, moderates the link between health-risk appraisal and protective behavioral intention. An empirical study targeting case city residents is underway. Findings advance IS theory on ethical AI design, citizen-consumer behavior, and AX transformation across public and commercial digital service platforms.

KEY WORDS

algorithm aversion, algorithm appreciation, protection motivation theory, ethical AI, citizen-consumer behavior, AX transformation

C1

An Exploratory Framework for Flagging Greenwashing: Combining NLP-Based Patent Analysis with Financial-Patent Efficiency Metrics

Wen Yan Toh, Wenting Liu (Singapore University of Social Sciences) and
Nat Liang (Independent Researcher)

As AI transforms e-commerce marketing, concerns about corporate greenwashing have intensified. Existing detection approaches largely rely on self-reported disclosures, which are subject to the biases they aim to assess. This study proposes a multi-method framework integrating NLP-based analysis of green patent texts with financial and patent efficiency metrics. Applying the framework to 76 U.S. publicly listed firms and 800 green patents across four sectors, we combine correlation analysis of R&D expenditure and patent output, NLP techniques including sentiment analysis, keyword extraction, and LDA topic modelling, and sector-level efficiency comparisons. The results reveal weak relationships between R&D spending and green patent output, substantial cross-sector variation in innovation efficiency, and systematic positivity bias in patent language amplified by editorial standardisation. These patterns are consistent with input-output misalignment and narrative amplification characteristic of symbolic environmental practices. The framework offers a scalable approach to screening sustainability claims for e-commerce platforms, ESG evaluation, and regulatory monitoring.

KEY WORDS

Greenwashing, Natural Language Processing, Green Patents, Ethical AI, Sentiment Analysis, Corporate Sustainability

An Exploratory Framework for Flagging Greenwashing: Combining NLP-Based Patent Analysis with Financial-Patent Efficiency Metrics

Toh W. Y. [a], *Liu, W.*[*a] *and Liang N.* [b]
[a] *School of Business, Singapore University of Social Sciences*
E-mail: wytoh003@suss.edu.sg; wentingliu@suss.edu.sg
[b] *Independent Researcher*
E-mail: milk.socoldnaty@gmail.com

Abstract

As AI transforms e-commerce marketing, concerns about corporate greenwashing have intensified. Existing detection approaches largely rely on self-reported disclosures, which are subject to the biases they aim to assess. This study proposes a multi-method framework integrating NLP-based analysis of green patent texts with financial and patent efficiency metrics. Applying the framework to 76 U.S. publicly listed firms and 800 green patents across four sectors, we combine correlation analysis of R&D expenditure and patent output, NLP techniques including sentiment analysis, keyword extraction, and LDA topic modelling, and sector-level efficiency comparisons. The results reveal weak relationships between R&D spending and green patent output, substantial cross-sector variation in innovation efficiency, and systematic positivity bias in patent language amplified by editorial standardisation. These patterns are consistent with input–output misalignment and narrative amplification characteristic of symbolic environmental practices. The framework offers a scalable approach to screening sustainability claims for e-commerce platforms, ESG evaluation, and regulatory monitoring.

Keywords: Greenwashing, Natural Language Processing, Green Patents, Ethical AI, Sentiment Analysis, Corporate Sustainability

Introduction

The rapid integration of artificial intelligence (AI) into electronic commerce has fundamentally reshaped how businesses communicate sustainability commitments to consumers. AI-powered marketing, personalised recommendation systems, and digital supply chain tools have enabled unprecedented reach and persuasiveness in corporate environmental messaging. However, this transformation has also raised concerns about greenwashing — the practice of making misleading claims about the environmental benefits of products, services, or corporate strategies (Lyon and Maxwell 2011). As digital platforms become primary interfaces between firms and consumers, the volume and complexity of sustainability-related messaging have outpaced the ability of stakeholders to systematically verify such claims.

While a substantial body of literature has examined greenwashing through corporate disclosures, ESG ratings, and consumer surveys, most existing approaches rely on self-reported data. While informative, these approaches are inherently constrained by their dependence on firm-generated or perception-based data, which may reflect the same biases they seek to detect. This limitation highlights the need for complementary, data-driven approaches that assess environmental claims against more objective indicators of technological and innovative activity.

Recent advances in natural language processing (NLP) offer new opportunities to analyse large-scale textual data in a systematic and replicable manner. In particular, NLP techniques can be applied to technical and semi-structured documents, such as patent abstracts, to extract thematic, linguistic, and sentiment-based signals of innovation activity. When combined with quantitative indicators such as R&D expenditure and patent output, these methods enable a multi-dimensional assessment of the relationship between reported environmental effort and measurable innovation outcomes.

The research is guided by two exploratory questions: (1) What patterns emerge in the association between R&D expenditure, green patent output, and financial performance across sectors, and do these patterns align with what the greenwashing literature would predict? (2) What linguistic and thematic characteristics of green patent filings may signal symbolic rather than substantive environmental innovation? Specifically, this study proposes an exploratory analytical framework that integrates three components: (1) correlation analysis examining the association between R&D expenditure, green patent output, and financial performance; (2) sector-level efficiency analysis comparing patent output relative to R&D inputs and (3) NLP-based analysis of green patent texts using keyword extraction, Latent Dirichlet Allocation (LDA) topic modelling, and sentiment analysis. Using a dataset of 76 U.S. publicly listed firms and 800 green patents, the framework aims to identify patterns of alignment or misalignment across inputs, outputs, and communication signals.

The study does not seek to directly identify instances of greenwashing. Rather, it focuses on detecting structural patterns, such as weak associations between R&D investment and patent output, sector-level disparities in innovation efficiency, and systematic linguistic framing, which are consistent with symbolic rather than substantive environmental activity. These patterns provide a scalable basis for further investigation and can inform the development of AI-enabled transparency tools for digital commerce environments, ESG evaluation systems, and regulatory monitoring. This study contributes to the theme of ethical challenges in AI commerce by demonstrating how AI and NLP techniques can serve as transparency tools within the e-commerce ecosystem, enabling stakeholders to evaluate the authenticity of corporate environmental claims made through digital channels.

Literature Review

2.1 Dimensions and Mechanisms of Greenwashing

Greenwashing, the practice of making misleading claims about environmental benefits, manifests through three principal mechanisms. The first is selective disclosure, whereby firms strategically highlight favourable environmental actions while concealing negative ones, enabling them to present curated sustainability narratives that obscure the full scope of their environmental impact (Lyon and Maxwell 2011). The second is symbolic corporate environmentalism, involving superficial gestures such as green marketing campaigns or token policy changes that project environmental responsibility without substantive operational transformation (Lyon and Montgomery 2015). The third is deceptive marketing, which exploits vague or unsubstantiated claims and imagery to exaggerate environmental benefits, leveraging consumer psychology and trust to create misleading perceptions (Lyon and Montgomery 2015). All three mechanisms are sustained by information asymmetry between corporations and their stakeholders, compounded by the absence of standardised reporting frameworks that would enable external verification. These mechanisms are difficult to detect using conventional approaches because they operate precisely within the gap between what firms say and what they do. As the primary evidence base in most greenwashing studies, self-reported sustainability data is itself shaped by the selective disclosure it seeks to investigate. This circularity motivates this study's focus on green patent a less easily manipulated indicator of green innovation output.

2.2 Motivations and Consequences of Greenwashing

Firms engage in greenwashing for interrelated strategic reasons. Legitimacy theory posits that organisations use voluntary sustainability reporting to ensure their activities are perceived as socially acceptable (Suchman 1995). Zharfpeykan (2021) provides empirical support for this in the mining/metals and financial services sectors, where companies selectively emphasise positive sustainability indicators while downplaying negative impacts to manage reputational risk. Greenwashing also serves investor-facing objectives. Cohen et al. (2023) find that firms with strong

ESG profiles attract more institutional investors, creating incentives to signal environmental commitment, including through the adoption of ESG-linked executive compensation, even when underlying practices are not fully aligned. This signalling behaviour is particularly relevant to the e-commerce context, where AI-powered marketing amplifies the reach and persuasiveness of environmental messaging across digital channels.

The consequences of greenwashing extend beyond reputational harm. Chen and Chang (2013) demonstrate that greenwashing erodes green trust by creating consumer confusion and increasing perceived risk, making consumers less likely to trust any environmental claims, including legitimate ones. At the market level, Lyon and Montgomery (2015) find a negative association between greenwashing exposure and cumulative abnormal returns, indicating tangible financial penalties when deceptive practices are uncovered. These findings suggest that green innovation activities and financial performance may be decoupled in firms engaged in symbolic environmentalism, which is a pattern this study explores at the cross-sectional level.

2.3 Industry Variations and the Quantity-Quality Gap

Greenwashing practices differ substantially across sectors. Zharfpeykan (2021) finds that mining companies in Australia tend to produce more balanced environmental reporting than financial services firms, likely because direct environmental impact invites greater scrutiny. Ling et al. (2024), examining the relationship between green bond issuance and ESG greenwashing among Chinese listed companies, find that green bonds can mitigate greenwashing tendencies by enhancing information transparency and alleviating financing constraints, though this effect varies across ESG dimensions and firm characteristics. Together, these findings suggest that the effectiveness of green signalling mechanisms depends heavily on sectoral context, regulatory exposure, and firm-level conditions, raising concerns that easily quantifiable metrics such as bond issuance or patent counts may create a facade of environmental responsibility without ensuring substantive change.

These differences have implications for how R&D expenditure should be interpreted as an indicator of environmental commitment. In sectors where green innovation is central to operations (e.g., energy, materials), one might expect a closer correspondence between R&D spending and green patent output. In sectors where environmental innovation is peripheral or where sustainability serves primarily as a reputational tool, this correspondence may be weaker. Marquis et al. (2016) describe this latter pattern as executional greenwashing. The practice of publicising environmental investments without delivering proportionate results. This study explores whether such input-output misalignment is observable at the sector level by comparing patent efficiency (green patents per $1 billion R&D) across industries, treating notable disparities as preliminary signals warranting further investigation.

2.4 Regulatory Gaps and the Case for AI-Based Screening

Current regulatory frameworks face significant limitations in addressing greenwashing. The absence of universally accepted definitions for "green" activities creates opportunities for selective disclosure (Lyon and Maxwell 2011), while inconsistencies across jurisdictions enable regulatory arbitrage, particularly in rapidly expanding markets such as green bonds. Existing regulations are largely reactive, lagging behind evolving greenwashing tactics. These gaps underscore the need for scalable, automated screening approaches. NLP techniques including sentiment analysis, topic modelling, and keyword extraction can systematically analyse corporate communications at a volume and speed that manual review cannot match. This capability is particularly relevant in AI-driven e-commerce ecosystems, where sustainability claims proliferate across digital channels. However, most existing NLP applications in this domain analyse self-reported corporate texts such as sustainability reports and press releases. This study instead applies NLP to patent abstracts that describe tangible technical innovations and are subject to legal examination, offering a less easily manipulated textual basis for assessing how firms frame their environmental innovation. In addition, the comparison between original and DWPI-standardised patent texts further provides a novel methodological lens. DWPI rewrites are produced by domain experts to improve technical clarity, standardise terminology, and enhance searchability. Yet the sentiment analysis in this study reveals that this clarification process incidentally shifts the sentiment distribution toward greater positivity and reduced neutrality. This is a noteworthy finding for researchers who rely on DWPI texts for NLP-based analysis: editorial standardisation intended to

improve readability may introduce a systematic tonal shift that, if unaccounted for, could inflate assessments of how positively firms frame their environmental innovations. The effect of data preprocessing choices on greenwashing detection outcomes has not been previously examined in the literature.

Methodology

3.1 Research Design Overview

The framework is implemented through three sequential stages, each producing outputs that feed into the next. Stage 1 (Financial Data Collection) assembles firm-level financial variables for 76 U.S. publicly listed companies and computes correlation coefficients between R&D expenditure, green patent counts, and financial performance indicators. The output of this stage is a cross-sectional mapping of input–output relationships at the aggregate and sector levels, including patent efficiency ratios (green patents per $1 billion R&D). Stage 2 (Patent Data Collection) compiles a corpus of 800 green patent texts from the Derwent World Patents Index, extracting both original and DWPI-standardised abstracts alongside structured metadata. This corpus serves as the input for Stage 3 (NLP Analysis), which applies three complementary techniques: keyword extraction and frequency analysis to characterise sector-level vocabulary profiles, Latent Dirichlet Allocation topic modelling to identify latent thematic structures, and lexicon-based sentiment analysis to assess tonal framing across both original and standardised texts. The convergence of findings across all three stages forms the basis for identifying patterns consistent with symbolic rather than substantive environmental innovation.

3.2 Financial Data Collection

This study adopted a sequential design integrating quantitative financial and patent data analysis with qualitative NLP techniques. This approach follows established precedent in greenwashing research, where triangulation between financial indicators and textual analysis strengthens the validity of findings (Lyon and Montgomery 2015; Zharfpeykan 2021). The quantitative phase examined the relationship between R&D expenditure, green patent output, and firm-level financial performance, while the qualitative phase applied NLP to uncover linguistic patterns indicative of symbolic environmentalism in patent filings. The sample comprised 76 publicly listed U.S. companies, selected using three criteria: (1) active listing on a major U.S. stock exchange to ensure audited financial data availability; (2) complete financial records for the fiscal years 2022–2024, minimising attrition from missing data; and (3) ranking among the top patent holders within their respective Global Industry Classification Standard (GICS) sectors for green technology filings. This purposive sampling strategy targeted firms most likely to exhibit measurable green innovation activity, enabling meaningful cross-sector comparison. Financial variables extracted included R&D expenses (RDE) as the primary independent variable, alongside control variables: firm size (SIZE, natural logarithm of total assets), leverage (LEV, total liabilities over total assets), return on assets (ROA), return on equity (ROE), Tobin's Q as a market valuation proxy, growth (retained earnings over net assets), and free cash flow (FCF, net operating cash flow over total assets). All variables were standardised to mitigate scale differences, and logarithmic transformations were applied where distributions were heavily skewed.

3.3 Patent Data Collection

To assess green innovation output, a dataset of 800 newly granted green patents was compiled, 200 per industry sector (Consumer Goods, Fashion, Manufacturing, and Petrochemical), covering the 2023–2024 grant period. This sample size was chosen to balance analytical depth with cross-sector comparability while ensuring sufficient textual material for NLP processing. Patents were sourced from the Derwent World Patents Index (DWPI), which provides professionally revised titles and abstracts authored by domain experts. The DWPI texts standardise terminology, reduce jargon, and fill gaps in original filings, making them particularly suited for automated textual analysis and cross-industry comparison. Crucially, the availability of both original and DWPI-revised texts enabled a comparative sentiment analysis that formed a novel component of this study. Key fields extracted included original and DWPI-revised titles and abstracts, assignee and ultimate parent information, application and publication years, and the count of cited patent references—the latter serving as a proxy for patent quality and technological embeddedness. These fields allow for multi-layered analysis of both the

content and structure of green patents enabling textual analysis (e.g., topic modelling, sentiment analysis) as well as quantitative examination (e.g., patent counts by firm, patent efficiency metrics, assignee-level innovation mapping).

3.4 NLP Techniques

Three complementary NLP techniques were applied to the patent corpus. First, keyword extraction and frequency analysis identified dominant terms within each sector to assess whether patent language reflects environmental substance or generic technical development. Second, Latent Dirichlet Allocation (LDA) topic modelling uncovered latent thematic structures, with the number of topics optimised using coherence scores. Third, lexicon-based sentiment analysis classified patent texts as positive, negative, or neutral for both original and DWPI-standardised versions, enabling a paired comparison of whether editorial standardisation shifts tonal distribution. All preprocessing was conducted using NLTK and spaCy in Python.

Several validation steps were undertaken to strengthen robustness. LDA coherence scores (C_v) were computed across topic numbers k = 3 to 10, and the optimal model was re-estimated across five random initialisations to confirm stability. The sentiment classifier was validated against 100 manually annotated abstracts (25 per sector) coded independently by two reviewers, achieving inter-rater agreement (Cohen's kappa) above 0.75 and classifier-to-human agreement above 80%. Keyword results were cross-checked against TF-IDF rankings to verify sector-distinctiveness. These procedures provide a reasonable basis for confidence in the reported patterns within the scope of this exploratory study.

Results

4.1 Quantitative Financial Analysis

The correlation between R&D expense (RDE) and applicant patent count is slightly negative (−0.03), while its correlation with parent patent count is near zero (0.04). Both coefficients are close to zero, indicating no meaningful linear association between R&D spending and green patenting activity. Similarly, green patent counts exhibit negligible relationships with key financial performance metrics. As shown in the correlation matrix (Figure 1), correlations between patent counts (both applicant and parent level) and Return on Assets (ROA) and Return on Equity (ROE) range from −0.04 to 0.04. These values suggest that variation in green patent output is not associated with short-term profitability or operational efficiency. Taken together, the results indicate a decoupling between (i) R&D inputs and patent-based innovation outputs, and (ii) patent outputs and firm-level financial performance. Firms with higher R&D spending do not systematically produce more green patents, and firms with greater green patent counts do not exhibit stronger financial performance within the observed period. Several mechanisms may explain these patterns. R&D expenditure may be allocated to activities that do not result in patentable outputs, including process improvements, regulatory compliance, or customer-focused innovation. In addition, firms may differ in their patenting strategies, with some relying on trade secrecy or other non-patent forms of knowledge protection. From a strategic perspective, green patenting may also be influenced by longer-term positioning, regulatory alignment, or signalling considerations rather than immediate financial returns. The weak relationship observed here is consistent with prior findings that inputs and outputs of environmental innovation are not always aligned (Lyon & Maxwell, 2011; Marquis et al. 2016).

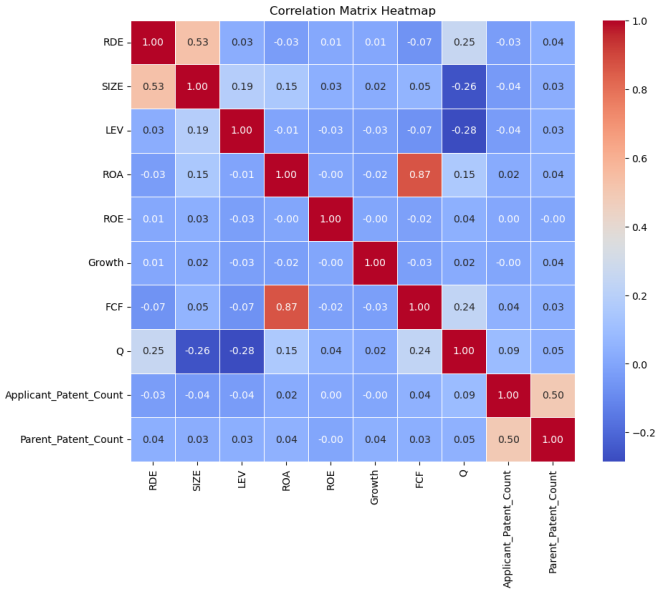


Figure 1. Correlation Matrix of R&D Expenditure, Green Patent Counts, and Financial Performance Metrics

4.2 Sector-Level Patent Efficiency

The results in Table 1 reveal substantial disparities across sectors. The Energy sector exhibits the highest efficiency, generating approximately 22.6 green patents per $1 billion of R&D expenditure, followed by Communication Services (15.76) and Consumer Staples (15.00). In contrast, Health Care produced minimum green patents despite notable R&D investment, while Industrials, the highest R&D spender overall—generated only 3.56 patents per $1 billion. Information Technology and Materials show moderate efficiency levels (8.94 and 7.92, respectively).

Table 1. Patent Efficiency by Sector (Patents per $1B R&D)

Sector	Patents/$1B R&D
Energy	22.60
Communication Services	15.76
Consumer Staples	15.00
Information Technology	8.94
Materials	7.92
Industrials	3.56
Health Care	0.01

These findings extend the aggregate results by showing that the weak relationship between R&D spending and green patent output is not uniform across sectors. Instead, substantial heterogeneity exists in how effectively R&D inputs are translated into patentable environmental innovation. In particular, high-efficiency sectors, such as Energy, are likely to reflect regulation-driven innovation. Firms in this sector operate under intense environmental scrutiny and compliance requirements, which incentivise targeted, patentable green technologies. Under such conditions, R&D investment is more directly channelled into measurable environmental outputs, resulting in high patent efficiency. Similarly, Communication Services and Consumer Staples may benefit from low-capital, high-scalability innovation structures, where incremental technological improvements—such as energy-efficient infrastructure or sustainable packaging—are more easily codified and patented. These sectors may also face strong consumer-facing reputational pressures, encouraging visible and countable forms of environmental innovation.

In contrast, low-efficiency sectors such as Industrials and Materials exhibit characteristics of capital-intensive and process-oriented innovation. R&D in these sectors is often directed toward large-scale operational improvements, supply chain optimisation, or incremental efficiency gains that are not always patentable or classified as "green." As a result, substantial R&D spending does not translate into proportional patent output. The Health Care sector represents a distinct case, where R&D is heavily concentrated in pharmaceutical and biomedical innovation, domains that are largely orthogonal to conventional green patent classifications. The absence of green patents despite significant R&D investment therefore reflects sectoral misalignment in classification frameworks, rather than necessarily a lack of innovation activity. Overall, these findings indicate that green patent efficiency is shaped not only by the scale of R&D investment, but also by sector-specific incentives, appropriability strategies, and the visibility of innovation outputs. Sectors with strong regulatory pressure and reputational exposure tend to produce more patentable and observable green outputs, while sectors with less visible or non-patentable innovation pathways contribute to the weak aggregate relationships observed earlier.

4.3 Keyword and Topic Modelling Results

Keyword distributions (Figure 2) reveal that green patent abstracts are dominated by general-purpose technical language — terms such as system, method, and data — rather than environmentally specific terminology, suggesting that many patents classified as 'green' are embedded within broader technological development. Sectoral vocabularies reinforce this pattern: Consumer Goods and Fashion are characterised by digital and design-oriented terms (user, interface, display), Manufacturing by production-centric language (material, process, metal), and Petrochemicals by fossil-based terminology (hydrocarbon, oil, feedstock), with environmental elements appearing only as incremental additions. The consistently low semantic specificity of environmental claims across sectors suggests that keyword-level classification may overstate the extent of sustainability-oriented innovation, reinforcing the interpretation that patent counts alone are a weak proxy for substantive environmental activity

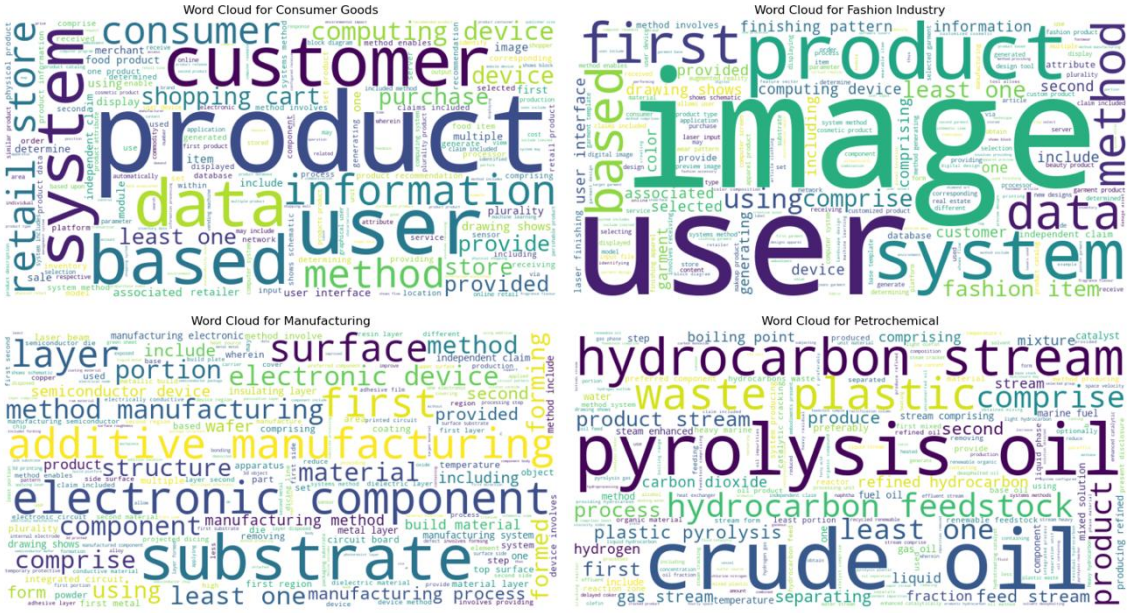


Figure 2. Wordcloud of Keyword Analysis

LDA topic modelling shows that dominant themes are structured around commercial and operational functions rather than environmental outcomes. In Consumer Goods and Fashion, topics centre on digital platforms and design processes, indicating that green-labelled patents are embedded in consumer engagement and product development rather than core environmental technologies. Manufacturing is dominated by process optimisation and materials engineering, where environmental benefits are indirect and not explicitly articulated. Petrochemicals exhibit a dual structure, with environmentally framed processes (e.g., pyrolysis, waste) layered onto dominant fossil-based themes (e.g., crude, hydrocarbon), suggesting incremental adaptation rather than structural transition. Across sectors, the persistence of generic constructs (e.g., method, device) indicates low thematic specificity, limiting interpretability.

Overall, topic structures suggest that patent-based signals of green innovation are broadly defined and weakly differentiated and may overstate substantive environmental transformation.

4.4 Sentiment Analysis

Sentiment analysis reveals systematic positivity bias across all sectors, with positive sentiment dominating in Consumer Goods (177/200) and Fashion (171/200), while Manufacturing and Petrochemicals exhibit greater tonal variation and higher negative counts (e.g., 80 in Petrochemicals). This divergence suggests that consumer-facing sectors emphasise perception-oriented framing, whereas regulated industries adopt more cautious signalling. Comparative analysis between original and DWPI-standardised texts reveals that standardisation substantially reduces neutral sentiment (e.g., from 35 to 5 in Consumer Goods, from 48 to 6 in Fashion) while elevating positive sentiment, replacing tentative language with more assertive expressions. This editorial amplification may unintentionally reinforce the appearance of environmental commitment in green patent narratives.

Conclusion

The findings reveal consistent misalignment between R&D inputs, patent outputs, and financial outcomes, suggesting that green innovation signals may not reflect substantive environmental progress. NLP analysis further shows that green patent narratives are vague and systematically positively framed, with editorial standardisation amplifying this effect. These patterns indicate that patent-based indicators function not only as signals of innovation but also as instruments of narrative construction — a concern amplified in AI-driven e-commerce environments where sustainability claims are disseminated at scale. Methodologically, the study demonstrates that combining patent analytics with NLP provides a scalable framework for screening sustainability claims applicable to ESG evaluation, regulatory monitoring, and e-commerce platforms. Several limitations constrain these findings. The study relies on patent data as a proxy for innovation, which may understate activity in sectors favouring trade secrecy or process-based improvements. The cross-sectional design captures associations at a single point, and the sample is limited to U.S. publicly listed firms across four sectors, leaving the framework's generalisability to other industries and jurisdictions untested. The lexicon-based sentiment classifier, while validated against a manually annotated subset, does not capture domain-specific connotations. Future research should incorporate longitudinal analysis, extend the framework to additional sectors and non-U.S. contexts, and explore transformer-based sentiment models fine-tuned on patent corpora

References

Chen, Y. S., & Chang, C. H. 2013. “Greenwash and green trust: The mediation effects of green consumer confusion and green perceived risk”. *Journal of business ethics*, *114*(3), 489-500.

Cohen, S., Kadach, I., Ormazabal, G., and Reichelstein, S. 2023. “Executive Compensation Tied to ESG Performance: International Evidence,” *Journal of Accounting Research* (61:3), pp. 789–844.

Lyon, T. P., and Maxwell, J. W. 2011. “Greenwash: Corporate Environmental Disclosure under Threat of Audit,” *Journal of Economics & Management Strategy* (20:1), pp. 3–41.

Lyon, T. P., and Montgomery, A. W.2015. “The Means and End of Greenwash,” *Organization & Environment* (28:2), pp. 223–249.

Marquis, C., Toffel, M. W., & Zhou, Y. 2016. “Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing,” *Organization science* (*27*:2), 483-504.

Ling, F., Zhen, S., Liang, C., & Zhao, Z. 2024. “Green bonds and corporate environmental, social, and governance greenwashing: evidence from China,” *The Singapore Economic Review*, 1-29.

Suchman, M. C. 1995. “Managing Legitimacy: Strategic and Institutional Approaches,” *Academy of Management Review* (20:3), pp. 571–610.

Zharfpeykan, R. 2021. “Representative Account or Greenwashing? Voluntary Sustainability Reports in Australia’s Mining/Metals and Financial Services Industries,” *Business Strategy and the Environment* (30:4), pp. 2209-2223.

KITS·ICEC·KIISS
2026 통합 국제학술대회

C2 Digital Media, Advertising & Content

Session Chair : Jee-Hae Lim
(Univ. of Hawaii at Mānoa)



C2

Who Tells It Better? Aligning Creation Modalities with Types of PSAs to Enhance Public Attitudes via Systematic Thinking

Xiaoxiao Gong, Danfeng Yuan (GuiZhou university)

Despite the significant advancements and potential of AI in public service announcements (PSAs), its current state of development suggests that it has not yet fully realized its promise, particularly in terms of public attitudes to AI-generated PSAs. This study investigates the impact of creation modalities (AI-generated vs. human-created) on public attitudes toward PSAs. Utilizing a mixed-methods approach with a sentiment analysis pretest and four experiments, we find that human-created PSAs generally elicit more positive public attitudes than AI-generated PSAs. Mediation analysis confirms that systematic thinking mediates this relationship. Crucially, the effect of creation modalities is moderated by the types of PSAs: for factual PSAs, human-created modalities lead to higher systematic thinking and more positive attitudes; conversely, for narrative PSAs, AI-generated modalities are more effective in triggering systematic thinking and positive attitudes. Furthermore, these interaction effects are moderated by audience involvement, with high-involvement audiences showing greater sensitivity to the creation modalities. These findings clarify the mechanism and boundary conditions of AI application in PSAs, providing evidence for matching creation modalities with specific types of PSAs.

KEY WORDS

Artificial intelligence, Public service announcements, PSAs creation modalities

C2

Can Machines Read Your Face? A Video Analytics Framework for Measuring Empathy via Emotional Mimicry in Video Recordings

Li Cui (The Hong Kong Polytechnic University), Jee-Hae Lim (University of Hawaii, Manoa),
Zhoudao Lu (University of Oxford),
Ka Chung Ng, Jingran Zhao (The Hong Kong Polytechnic University)

Empathy is a key driver of individual and organizational success, and recent advances in machine learning enable scalable extraction of such social signals from digital data. We propose a video analytics framework to measure empathy through an Affective Mimicry Index (AMI), capturing emotional mimicry–the alignment of expressed affect between individuals–as a behavioral proxy for empathy. Using televised CEO interviews, we compute AMI by quantifying mimicry between CEOs and interviewers across question–answer interactions. We then examine its association with firm outcomes. Results show that higher CEO AMI is linked to improved workplace safety, higher firm value, and fewer regulatory penalties. These findings suggest that empathic leadership enhances employee welfare and organizational performance. Methodologically, this study integrates computer vision and conversational analytics to measure empathy from video data, offering new insights for business analytics and leadership research.

KEY WORDS

Video analytics, FinTech, empathy, computational social science

Can Machines Read Your Face? A Video Analytics Framework for Measuring Empathy via Emotional Mimicry in Video Recordings

Cui, Li[a] and Lim, Jee-Hae[b] and Lu, Zhoudao[c] and Ng, Ka Chung[d] and Zhao, Jingran[e]
[a] School of Accounting and Finance, The Hong Kong Polytechnic University
E-mail: li2049.cui@connect.polyu.hk
[b] Shidler College of Business, University of Hawai'i at Mānoa
E-mail: jeehae@hawaii.edu
[c] Department of Statistics, University of Oxford
E-mail: zhoudao.lu@kellogg.ox.ac.uk
[d] Department of Management and Marketing, The Hong Kong Polytechnic University
E-mail: kc-boris.ng@polyu.edu.hk
[e] School of Accounting and Finance, The Hong Kong Polytechnic University
E-mail: jingran.zhao@polyu.edu.hk

Abstract

Empathy is a key driver of individual and organizational success, and recent advances in machine learning enable scalable extraction of such social signals from digital data. We propose a video analytics framework to measure empathy through an Affective Mimicry Index (AMI), capturing emotional mimicry—the alignment of expressed affect between individuals—as a behavioral proxy for empathy. Using televised CEO interviews, we compute AMI by quantifying mimicry between CEOs and interviewers across question–answer interactions. We then examine its association with firm outcomes. Results show that higher CEO AMI is linked to improved workplace safety, higher firm value, and fewer regulatory penalties. These findings suggest that empathic leadership enhances employee welfare and organizational performance. Methodologically, this study integrates computer vision and conversational analytics to measure empathy from video data, offering new insights for business analytics and leadership research.

Keywords: Video analytics, FinTech, empathy, computational social science

Introduction

Human emotional and social abilities—such as empathy, conflict resolution, and emotional regulation—play a key role in shaping individual and organizational outcomes (Ou et al., 2024). Traditionally, these behaviors are measured through surveys, lab experiments, and manual coding, which are costly, time-consuming, and hard to scale. As a result, their use in large, real-world settings is limited. Recent advances in machine learning offer scalable alternatives by extracting social signals from digital data. For example, natural language processing can infer strategic communication, computer vision can assess managerial trustworthiness, and multimodal analytics can measure employee creativity—highlighting the potential of data-driven approaches to operationalize complex human behaviors at scale..

In this paper, we propose a video analytics framework to measure emotional mimicry—an essential social behavior that facilitates empathy—using interview recordings. Emotional mimicry refers to the unconscious alignment of one individual's emotional expressions with another's and is widely

recognized as a core mechanism underlying empathy (Hess et al., 2014). Neuroscience and psychology research suggest that individuals with higher empathy exhibit stronger mimicry responses, making it a reliable behavioral proxy for empathy. Building on this literature, we develop a machine learning framework to construct an Affective Mimicry Index (AMI) from interview videos. Specifically, we segment interviews into question–answer (QA) pairs, track facial expressions of both speakers, and quantify emotional alignment over time. We then aggregate these measures to obtain a CEO-level AMI.

Capturing empathy is important for two main reasons. First, empathy is fundamental to interpersonal relationships. It enables individuals to understand and resonate with others' emotional states, facilitating cooperation, communication, and social bonding. Second, empathy is a key component of leadership effectiveness. Prior research shows that managerial traits significantly influence corporate decisions and performance, and empathy has increasingly been recognized as a critical leadership capability.

To demonstrate the feasibility and value of our approach, we analyze 1,526 CEO television interview videos from CNBC (2013–2017), covering 500 CEOs. We construct AMI by quantifying emotional mimicry between CEOs and interviewers within QA segments. Our framework integrates large language models for QA segmentation, conversational analytics for interaction structuring, and computer vision for emotion recognition. We validate the proposed index in three ways. First, AMI is positively associated with CEO ratings on Glassdoor, suggesting it captures perceived empathy. Second, AMI is negatively correlated with CEO narcissism proxies, consistent with psychological theory. Third, AMI aligns closely with human-annotated empathy ratings, indicating strong measurement validity.

We then examine the relationship between CEO empathy and firm outcomes. We find that higher AMI is associated with improved workplace safety, fewer regulatory penalties, and higher firm value (Tobin's Q). These results suggest that empathic CEOs are better at fostering positive workplace environments, managing crises, and enhancing organizational performance.

This study contributes to the literature in three ways. First, it advances research on machine learning-based measurement of human social behavior by proposing a scalable, theory-driven framework that integrates LLMs, conversational analytics, and computer vision. Second, it contributes to AI in finance by demonstrating how video analytics can extract firm-relevant behavioral signals from executive communication. Third, it deepens our understanding of how executive soft skills, particularly empathy, influence corporate outcomes, offering implications for leadership selection and organizational strategy.

Literature Review

Empathy in Business Research

Empathy is a key component of leadership and organizational effectiveness, shaping communication, trust, and decision-making (Clark et al., 2019; Ou et al., 2024). Empathic leaders foster cooperation, improve employee outcomes, and enhance firm performance (Sadri et al., 2011).

Despite its importance, research on empathy faces two challenges. First, empathy is conceptually complex and multidimensional, involving affective and cognitive components as well as trait and state dimensions (Cuff et al., 2016), which are often blurred in measurement (Clark et al., 2019). Second, measurement approaches are limited. Most studies rely on self-reports like the Interpersonal Reactivity Index (Davis, 1983), which are prone to bias, while behavioral and neuroscientific methods are more objective but hard to scale (Batson, 2009; Adamopoulos et al., 2018). As a result, measuring empathy in real-world organizational contexts and linking it to firm outcomes remains difficult, motivating scalable, behavior-based approaches (Clark et al., 2019).

Emotional Mimicry and Empathy

Emotional mimicry—the spontaneous alignment of emotional expressions—is a core mechanism underlying empathy (Hess et al., 2014). It facilitates emotional contagion, enabling individuals to share others' feelings and develop empathic understanding (Hoffman, 1985).

Theoretical and empirical work supports this link. Mimicry has been proposed as a pathway to empathy (Rogers, 1975) and is associated with greater empathic concern and prosocial behavior (Dimberg &

Thunberg, 2012). Neuroscientific evidence further shows that mirror neuron systems enable individuals to simulate others' emotions, providing a biological basis for this relationship (Singer & Lamm, 2009).

However, prior studies measure mimicry mainly in laboratory settings, limiting external validity and highlighting the need for scalable methods to capture it in real-world interactions.

Multimodal Data Processing and Applications

Advances in multimodal data processing provide new opportunities to address these challenges. Video analytics integrates computer vision, natural language processing, and audio analysis to extract behavioral signals from unstructured data (Yang et al., 2023; Banker et al., 2024).

In business research, multimodal approaches have been used to analyze managerial communication, investor behavior, and customer interactions. For example, combining text and audio improves market predictions (Yang et al., 2023), while visual cues such as facial expressions influence perceptions of trustworthiness and decision-making (Blankespoor et al., 2017; Banker et al., 2024).

However, existing studies primarily focus on observable traits such as sentiment or attractiveness, with limited attention to latent social constructs such as empathy. Moreover, few studies integrate multimodal signals to capture interactional behaviors in executive contexts.

Methodology

We propose a comprehensive video analytics framework to quantify emotional mimicry in interview settings. As illustrated in Figure 1, the framework consists of three main components: (1) Question–Answer (QA) pair segmentation, (2) facial emotion recognition, and (3) Affective Mimicry Index (AMI) construction. Each component is designed to progressively transform raw interview videos into structured emotional alignment measures between interviewers and interviewees.

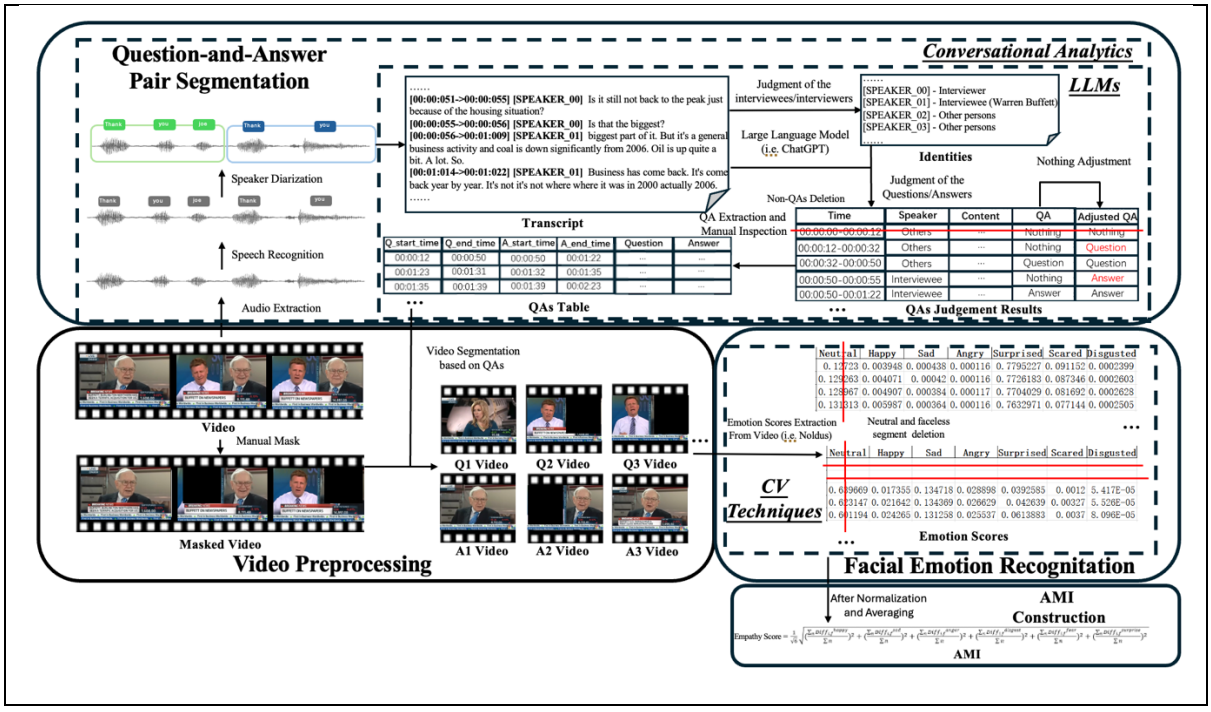


Figure 1. A Video-Based Framework for Measuring Emotional Mimicry

QA Pair Segmentation

The first stage focuses on isolating meaningful conversational units from raw interview videos. Since emotional mimicry is inherently tied to interpersonal interactions, accurately identifying question–

answer (QA) exchanges between interviewers and interviewees is essential. To this end, we rely on AI-based processing techniques. Rather than directly segmenting video streams, we adopt a two-step strategy: first constructing a QA pair table using audio and textual data, and then leveraging this table to segment the video.

As illustrated in Figure 1 (top-left and center), we begin by extracting audio from interview videos and performing speaker diarization using *Pyannote* (Bredin et al., 2019). This step identifies "who speaks when," allowing us to divide the audio into speaker-specific segments. We then apply *Whisper* (Radford et al., 2022) for automatic speech recognition (ASR), generating transcripts with precise timestamps. These transcripts are subsequently merged with diarization outputs to produce structured conversational data that include speaker identity, timing, and textual content.

A key challenge at this stage is identifying speaker roles, particularly distinguishing interviewers from the interviewee. Traditional NLP approaches, such as *GloVe* and *BERT*, are limited in their ability to capture complex conversational dynamics and infer roles from context. To address this limitation, we employ a large language model (LLM), specifically *GPT-4* (Achiam et al., 2023), which offers stronger reasoning capabilities and contextual understanding. Using carefully designed prompts, the model assigns each speaker to one of three categories—"Interviewer," "Interviewee," or "Other"—thereby enabling reliable role identification even in multi-speaker settings and ensuring clear differentiation between interviewers and the interviewee.

Building on this role identification, we further classify each sentence in the transcript as "Question," "Answer," or "NA" (non-relevant content, such as greetings, small talk, and closing remarks). This step is also implemented using *GPT-4* for sentence-level classification, allowing us to retain primarily those QA exchanges that are closely related to firm-relevant content and thereby improve the precision of empathy measurement.

The classification outputs are then integrated with the processed transcripts to construct a preliminary QA table. Because LLM predictions may occasionally be imperfect (e.g., misclassifying relevant sentences as "NA"), we implement a post-processing procedure to refine the results. In particular, we ensure that only correctly matched QA pairs are retained while removing misclassified or unmatched segments. For example, isolated misclassifications within otherwise coherent QA sequences are corrected. Finally, all entries labeled as "NA" are removed, resulting in a clean QA pair table with precise temporal boundaries. This table serves as the foundation for subsequent video segmentation.

Facial Emotion Recognition

The second component extracts emotional signals from the segmented videos (Figure 3, right panel). We employ *FaceReader* from *Noldus* (Loijens & Krips, 2013), a deep learning-based facial analysis tool, to quantify emotional expressions at the frame level.

FaceReader detects faces using advanced computer vision techniques and models facial features via 468 key landmarks. These features are mapped to emotional states using neural networks trained on large annotated datasets grounded in the Facial Action Coding System (FACS).

For each video frame, the model produces seven emotion scores—neutral, happiness, sadness, anger, disgust, fear, and surprise—each ranging from 0 to 1. Because mimicry of neutral emotions lacks clear social significance and is difficult to identify reliably, we exclude neutral expressions from the analysis, consistent with prior psychological research (Hess & Fischer, 2013). To ensure comparability and eliminate scale-related distortions, we then normalize the remaining six emotion scores so that they sum to one.

Given the high frame rate (e.g., 24 frames per second), this process generates a dense time series of emotional data. Frames without detectable faces are discarded to avoid introducing noise into subsequent mimicry calculations. The resulting output is an emotion score matrix for each QA segment, capturing the dynamic emotional profiles of both the interviewer and the interviewee.

Affective Mimicry Index (AMI) Construction

The final component quantifies emotional mimicry based on the alignment between interviewer and interviewee emotions. For each QA pair j in video i, we compute the average emotion score for each emotion $e \in \{happy, sad, angry, disgust, fear, and\ surprise\}$ across all frames within the question (Q) and answer (A) segments. Specifically, $Score^{Q}_{i,j,e}$ denotes the interviewer i's emotion during the question j, and $Score^{A}_{i,j,e}$ denotes the interviewee i's emotion during the answer j. These aggregated values capture the emotional states of both parties within each interaction.

We then define emotional similarity as the absolute difference between the interviewer's and interviewee's emotional scores: $Diff_{i,j,e} = \left|Score^{A}_{i,j,e} - Score^{Q}_{i,j,e}\right|$. This formulation captures how closely the interviewee's emotional response aligns with that of the interviewer. Because answers temporally follow questions, this measure reflects emotional mimicry rather than coincidental similarity.

To obtain a video-level measure of emotional alignment, we aggregate emotional differences across all QA pairs and emotion categories: $EmoAlign_i = \sqrt{\frac{1}{6}\sum_e (\frac{\sum_j Diff_{i,j,e}}{\sum j})^2}$. This root-mean-square formulation ensures balanced contributions across emotion types while assigning greater weight to larger deviations. Lower values of $EmoAlign_i$ indicate stronger emotional alignment between the interviewer and interviewee.

Finally, we transform emotional alignment into a measure of emotional mimicry: $EmoMimic_i = 1 - EmoAlign_i$ indicate stronger emotional mimicry. To capture individual-level behavior, we further aggregate $EmoMimic_i$ across all interviews involving the same interviewee, resulting in the Affective Mimicry Index (AMI).

Experiments

Research Setting, Data, and Validation

We apply our Affective Mimicry Index (AMI) to CEO television interview videos, as emotional mimicry is context-dependent and often used to convey intentions in social interactions (Hess & Fischer, 2013). CEO interviews are a form of strategic communication, where leaders shape investor perceptions and build rapport, making them ideal for observing intention-driven mimicry. The standardized format and real-time interaction allow consistent measurement of CEOs' facial responses.

Our sample includes 3,480 CNBC interview videos from 2013–2017, with 1,524 video observations from 500 CEOs. We construct a behavioral measure of empathy (CEOAMI) by averaging video-level emotional mimicry across interviews and complement it with CEO- and firm-level characteristics.

We validate AMI as a proxy for CEO empathy using three approaches. First, CEOAMI is positively associated with Glassdoor CEO ratings, suggesting that more empathetic CEOs are perceived more favorably by employees. Second, CEOAMI is negatively correlated with narcissism, measured using pronoun usage in conference calls ($r = -0.21$, $p < 0.001$). Third, human validation shows that CEOs with higher AMI are perceived as significantly more empathetic, with high-AMI videos receiving substantially higher ratings than low-AMI ones (difference = 1.19, $p < 0.001$; $r = 0.703$). These results provide robust evidence that AMI captures meaningful variation in CEO empathy and aligns closely with behavioral and perceptual measures of empathy.

Empirical Results

We examine the impact of CEO empathy on three key firm outcomes: workplace safety, crisis management, and firm value. Workplace safety is measured using OSHA establishment-level injury data (2002–2011), crisis outcomes using regulatory penalties from Violation Tracker (2000–2023), and firm value using Tobin's Q (1992–2023). Together, these outcomes capture distinct but related channels through which empathetic leadership affects organizational performance.

We begin with workplace safety, which reflects firms' ability to protect employee well-being and manage operational risks. Prior research suggests that safety-related policies and investments are shaped by top executives (Cohn & Wardlaw, 2016). Empathetic CEOs are more attentive to employees'

emotional and physical conditions and are therefore more likely to detect early warning signals and allocate resources toward preventive safety measures. In addition, by fostering trust and reducing workplace stress, empathetic leaders can improve employees' compliance with safety practices.

Table 1. CEO Empathy and Workplace Safety, crisis management, and Firm Value

Variable	*Injury/Hour*			*Ln(Penalty)*			*TobinQ*		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
CEOAMI	-0.073 * (-1.90)	-0.073 ** (-2.16)	-0.123 *** (-3.30)	-0.093 ** (-3.14)	-0.078 ** (-2.51)	-0.068 * (-1.84)	0.053 ** (2.20)	0.049 ** (2.36)	0.056 *** (2.66)
EstablishmentChar	No	Yes	Yes	No	No	No	No	No	N0
FirmChar	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CEOChar	No	No	Yes	No	No	Yes	No	No	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.254	0.302	0.306	0.203	0.211	0.213	0.151	0.401	0.419
Observations	11,972			3,447			4,426		
Sample Period	2002 to 2011			2000 to 2023			1992 to 2023		

Notes: This table presents the results of the relationship between CEO empathy and workplace safety, penalties for facility misconduct and firm value. All continuous variables are standardized to have a mean of 0 and a variance of 1. The *t*-statistics, estimated using robust standard errors, are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

To test this prediction, we estimate the following specification:

$$Injury/Hour_{i,j,t} = \beta_0 + \beta_1 CEOAMI_i + \boldsymbol{EstablishmentChar_{j,t}\gamma} + \boldsymbol{FirmChar_{i,t-1}\delta} + \boldsymbol{CEOChar_i\omega} + IndustryFE + YearFE + \varepsilon_{i,j,t}$$

where $Injury/Hour_{i,j,t}$ denotes establishment-level injury rates, $CEOAMI_i$ measures CEO emotional mimicry, and $\boldsymbol{EstablishmentChar}$, $\boldsymbol{FirmChar}$ and $\boldsymbol{CEOChar}$ represent establishment-, firm-, and CEO-level controls, respectively. In Table 1, for simplicity, we group $\boldsymbol{IndustryFE}$, $\boldsymbol{YearFE}$, and the constant term under the label, $\boldsymbol{Fixed\ Effects}$.

We first focus on workplace safety. Columns (1)–(3) under *Injury/Hour* in Table 1 report the corresponding results. In Column (1), we include only the CEO AMI and fixed effects, without incorporating any control variables. In Column (2), we add establishment- and firm-level controls. Finally, in Column (3), we include CEO characteristic control variables. Across all specifications, the coefficient on $CEOAMI$ is negative and statistically significant. The coefficient of *CEOAMI* reported in Column (3) is significant and positive, suggesting that firms led by empathetic CEOs tend to have lower injury rates. Furthermore, the economic significance of CEO empathy on injury rates is notable. Specifically, the coefficient of -0.123 implies that for each standard deviation increase in CEO AMI scores, the firms they lead experience a 10.61% reduction in injury rates.

We next examine crisis management outcomes. Managing corporate misconduct requires the ability to process soft information and respond effectively under uncertainty (König et al., 2020). Due to information asymmetry within firms, CEOs rely on internal signals to identify and mitigate potential violations (Heese et al., 2020). Empathetic CEOs are better equipped to recognize such signals and to communicate effectively with stakeholders, thereby reducing the severity of misconduct outcomes.

We estimate:

$$Ln(Penalty_{i,j,t}) = \beta_0 + \beta_1 CEOAMI_i + \boldsymbol{FirmChar_{i,t-1}\delta} + \boldsymbol{CEOChar_i\omega} + IndustryFE + YearFE + \varepsilon_{i,j,t}$$

As shown in Table 1 under *Ln(Penalty)*, *CEOAMI* is negatively associated with regulatory penalties across all specifications. Economically, a one standard deviation increase in CEO empathy leads to a 6.4% reduction in penalty amounts. This finding indicates that empathetic CEOs are more effective in mitigating the financial consequences of corporate misconduct, consistent with their superior ability to detect risks and manage crises.

Finally, we examine firm value. CEO empathy can influence firm value through two primary channels: enhancing performance and reducing risk. Empathetic CEOs foster stronger employee relationships and a more cohesive organizational culture, which contribute to long-term performance (Edmans, 2012). At the same time, by improving workplace safety and reducing misconduct, they lower operational and reputational risks, which can reduce the cost of capital.

To test this, we estimate:

$$TobinQ_{i,t} = \beta_0 + \beta_1 CEOAMI_i + \boldsymbol{FirmChar_{i,t-1}\delta} + \boldsymbol{CEOChar_i\omega} + IndustryFE + YearFE + \varepsilon_{i,t}$$

The results in Table 1 under *TobinQ* show that *CEOAMI* is positively and significantly associated with firm value. This suggests that CEO empathy is not only beneficial for internal organizational outcomes but also translates into higher market valuation.

Conclusion

In this paper, we introduce a video analytics framework to construct the AMI, a measure of emotional mimicry that serves as a proxy for CEO empathy using interview video recordings. Our framework integrates LLMs, conversational analytics, and computer vision techniques to capture subtle facial expressions and emotional cues from CEOs' responses to interviewers' questions. These cues indicate emotional mimicry and provide a reliable behavioral proxy for empathy. To demonstrate the effectiveness and practical value of AMI, we first conduct the measure with a series of tests. We then examine its business implications by analyzing its relationship with corporate policies (workplace safety) and firm performance (crisis management and firm value).

References

Adamopoulos, P., Ghose, A., and Todri, V. 2018. "The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms," *Information Systems Research* (29:3), pp. 612–640.

Banker, R. D., Ding, H., Huang, R., and Li, X. 2024. "Market Reaction to CEOs' Dynamic Hemifacial Asymmetry of Expressions," *Management Science* (70:7), pp. 4847–4874.

Batson, C. D. 2009. "These Things Called Empathy: Eight Related but Distinct Phenomena," in *The Social Neuroscience of Empathy*, J. Decety and W. Ickes (eds.), pp. 3–15.

Blankespoor, E., Hendricks, B. E., and Miller, G. S. 2017. "Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows," *Journal of Accounting Research* (55:2), pp. 275–327.

Bredin, H., Yin, R., Coria, J. M., Gelly, G., Korshunov, P., Lavechin, M., Fustes, D., Titeux, H., Bouaziz, W., and Gill, M.-P. 2019. "pyannote.audio: Neural Building Blocks for Speaker Diarization," in *Proceedings of Interspeech 2019*, pp. 272–276.

Clark, M. A., Robertson, M. M., and Young, S. 2019. "I Feel Your Pain: A Critical Review of Organizational Research on Empathy," *Journal of Organizational Behavior* (40:2), pp. 166–192.

Cohn, J. B., and Wardlaw, M. I. 2016. "Financing Constraints and Workplace Safety," *The Journal of Finance* (71:5), pp. 2017–2058.

Cuff, B. M. P., Brown, S. J., Taylor, L., and Howat, D. J. 2016. "Empathy: A Review of the Concept," *Emotion Review* (8:2), pp. 144–153.

Davis, M. H. 1983. "Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach," *Journal of Personality and Social Psychology* (44:1), pp. 113–126.

Dimberg, U., and Thunberg, M. 2012. "Empathy, Emotional Contagion, and Rapid Facial Reactions to Angry and Happy Facial Expressions," *PsyCh Journal* (1:2), pp. 118–127.

Edmans, A. 2012. "The Link Between Job Satisfaction and Firm Value, With Implications for Corporate Social Responsibility," *Academy of Management Perspectives* (26:4), pp. 1–19.

Heese, J., and Pérez-Cavazos, G. 2020. "When the Boss Comes to Town: The Effects of Headquarters' Visits on Facility-Level Misconduct," *The Accounting Review* (95:6), pp. 235–261.

Hess, U., and Fischer, A. 2013. "Emotional Mimicry as Social Regulation," *Personality and Social Psychology Review* (17:2), pp. 142–157.

Hess, U., and Fischer, A. 2014. "Emotional Mimicry: Why and When We Mimic Emotions," *Social and Personality Psychology Compass* (8:2), pp. 45–57.

Hoffman, M. L. 1985. "Interaction of Affect and Cognition in Empathy," in *Emotions, Cognition, and Behavior*, R. B. Zajonc (ed.), pp. 103–131.

König, A., Graf-Vlachy, L., Bundy, J., and Little, L. M. 2020. "A Blessing and a Curse: How CEOs' Trait Empathy Affects Their Management of Organizational Crises," *Academy of Management Review* (45:1), pp. 130–153.

Loijens, L., and Krips, O. 2013. *FaceReader Methodology*. Noldus Information Technology.

OpenAI, Achiam, J., Adler, S., Ahmad, L., Aleman, F. L., Altenschmidt, J., Altman, S., Balaji, S., Bao, H., Bernadett-Shapiro, G., Bogdonoff, L., Campbell, R., Cann, A., Carlson, C., Chen, D., Chen, S., Chen, J., Cho, C., Chu, C., and Zheng, T. 2023. "GPT-4 Technical Report," *arXiv preprint arXiv:2303.08774*.

Ou, A. Y., Lu, Q., Li, X., Chi-Nien, C., and Chen, G. 2024. "CEO Humility and Corporate Social Irresponsibility: Evidence Based on a New Unobtrusive Measure," *Organization Science* (35:6), pp. 2016–2039.

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., and Sutskever, I. 2022. "Robust Speech Recognition via Large-Scale Weak Supervision," *arXiv preprint arXiv:2212.04356*.

Rogers, R. W. 1975. "A Protection Motivation Theory of Fear Appeals and Attitude Change," *The Journal of Psychology* (91:1), pp. 93.

Sadri, G., Weber, T. J., and Gentry, W. A. 2011. "Empathic Emotion and Leadership Performance: An Empirical Analysis Across 38 Countries," *The Leadership Quarterly* (22:5), pp. 818–830.

Singer, T., and Lamm, C. 2009. "The Social Neuroscience of Empathy," *Annals of the New York Academy of Sciences* (1156), pp. 81–96.

Yang, Y., Qin, Y., Fan, Y., and Zhang, Z. 2023. "Unlocking the Power of Voice for Financial Risk Prediction: A Theory-Driven Deep Learning Design Approach," *MIS Quarterly* (47:1), pp. 63–96.

C2

A Cost-Sensitive Information Grading Model Considering Asymmetric Misclassification Risks

Seojin Lee, Wonyoung Kim, Hyeseong Lee, Hangbae Chang (중앙대학교 경영경제대학 산업보안학과)

Information classification is vital for protecting assets, yet existing models often overlook the asymmetric risks of misclassification. Specifically, misclassifying high-level information as low-level results in critical security breaches. We addressed this gap by formulating information grading as a cost-sensitive multi-class classification task. We developed an automated pipeline integrating a cost-weighted loss function with a strategic document-level aggregation method. We validated the model using two proprietary corporate datasets labeled at sentence and document levels. The experimental results demonstrated that our framework effectively suppressed fatal downward misclassifications and significantly reduced total costs compared to the Logistic Regression baseline. These findings suggest that cost-sensitive learning can improve the reliability of automated information grading in enterprise security environments.

KEY WORDS

Information Classification, Cost-Sensitive Learning, Asymmetric Risk, Misclassification Cost

A Cost-Sensitive Information Grading Model Considering Asymmetric Misclassification Risks

Lee, S. J.[a], Choi, Y. J.[b], Kim, W. Y.[c], Lee, H. S.[d] and Chang, H. B.[e]
[a] Convergence Security, Graduate School, Chung-Ang University
E-mail: emily9821@cau.ac.kr
[b] Convergence Security, Graduate School, Chung-Ang University
E-mail: yjji9783@cau.ac.kr
[c] Convergence Security, Graduate School, Chung-Ang University
E-mail: wonyoungkim@cau.ac.kr
[d] Convergence Security, Graduate School, Chung-Ang University
E-mail: lee0329@cau.ac.kr
[e] Industrial Security, College of Business and Economics, Chung-Ang University
E-mail: hbchang@cau.ac.kr

Abstract

Information classification is vital for protecting assets, yet existing models often overlook the asymmetric risks of misclassification. Specifically, misclassifying high-level information as low-level results in critical security breaches. We addressed this gap by formulating information grading as a cost-sensitive multi-class classification task. We developed an automated pipeline integrating a cost-weighted loss function with a strategic document-level aggregation method. We validated the model using two proprietary corporate datasets labeled at sentence and document levels. The experimental results demonstrated that our framework effectively suppressed fatal downward misclassifications and significantly reduced total costs compared to the Logistic Regression baseline. These findings suggest that cost-sensitive learning can improve the reliability of automated information grading in enterprise security environments.

Keywords: Information Classification, Cost-Sensitive Learning, Asymmetric Risk, Misclassification Cost

Introduction

As digital technologies advance and cloud computing becomes more widespread, the volume of data handled by enterprises has grown rapidly. This trend has also made the protection of information assets more complex (Tari 2014). In response, international and domestic information security standards, including ISO/IEC 27001 and ISMS-P, emphasize the need for an information classification framework that ensures information is categorized according to its value and sensitivity, and that appropriate levels of protection are applied accordingly (ISMS 2025).

However, prior studies on information classification have mainly focused on improving the overall performance of machine learning models, with performance typically evaluated using accuracy as well as metrics such as precision, recall, and F1-score (Allam et al. 2025). These approaches typically assume equal misclassification costs across all classes. As a result, they do not adequately reflect the asymmetric risks present in real-world information security environments.

In practice, the consequences of misclassification are highly asymmetric. **Downward misclassification** refers to cases in which highly confidential information is assigned to a lower security grade. Such errors can directly increase the likelihood of serious security incidents, including data breaches. For example, the "National Public Data" breach in the United States in August 2024 exposed approximately 2.9 billion

sensitive personal records. The incident has been associated with insufficient encryption, weak access control, and failures in information classification (Toulas 2024). This case illustrates the severe risks of treating high-grade information as low-grade. This asymmetry suggests that the potential damage caused by under-protection far outweighs the operational cost associated with over-protection. While **upward misclassification** may result in operational inefficiency, we focus on mitigating the more consequential downward errors.

Accordingly, we frame information classification as a cost-sensitive multiclass classification problem rather than a task of optimizing accuracy alone. The main contributions of this study are as follows:

- We reconceptualize information classification as a cost-sensitive problem that reflects asymmetric security risks.
- We propose a classification framework that combines cost-weighted learning with a conservative document-level aggregation strategy to penalize downward misclassification.
- We show through experiments on real-world corporate data that the proposed approach reduces critical misclassification risk and overall misclassification cost while maintaining practical predictive performance.

The remainder of this paper is organized as follows. Section 2 reviews related work and theoretical background. Section 3 describes the proposed methodology. Section 4 presents the system implementation and experimental results using real-world corporate data. Finally, Section 5 concludes the paper and discusses future research directions.

Related Work

Information Classification and the Need for Automation Approaches

Information classification is a fundamental management practice for protecting an organization's information assets, aiming to apply differentiated levels of security control based on the sensitivity and importance of the information. International information security standards, such as ISO/IEC 27001 and NIST guidelines, require a systematic grading process based on the value of information and the potential impact of its leakage (ISO 2022; NIST 2023). With the exponential growth of corporate data, research on automated classification frameworks using natural language processing and machine learning has increased to overcome the limitations of manual classification. Tan et al. (2024) highlighted the challenge of identifying sensitive information within large-scale corporate document repositories, emphasizing the necessity of content-based automated classification. Furthermore, Rāts (2021) pointed out inconsistencies in manual labeling rules within Enterprise Content Management (ECM) systems and proposed a tunable ML-based document classification framework. Most recently, Ignaczak et al. (2026) proposed an NLP model framework that estimates the intrinsic value and sensitivity of information using text features.

Asymmetric Misclassification Risks in Information Security

Previous studies have significantly contributed to improving the "overall accuracy" of classification by enhancing text representation methods or algorithm architectures. However, a fundamental limitation exists: most approaches utilize standard loss functions, such as cross-entropy, which treat misclassifications across all classes with equal penalty. In the realm of information security, the consequences of misclassification are fundamentally asymmetric. **Downward misclassification**, in which high-grade confidential information is assigned to a lower grade, can lead to serious security breaches, such as data leakage and loss of corporate competitiveness. Conversely, **upward misclassification**, in which general information is assigned to a higher grade, may result in operational inefficiencies due to excessive access controls. In other words, misclassification costs act asymmetrically across classes, and conventional accuracy-driven classification methods cannot effectively control the downward misclassification of the highest-risk data.

Cost-Sensitive Learning for Information Classification

Cost-Sensitive Learning (CSL) has emerged as a compelling alternative to address the structural flaws of accuracy-driven models. CSL is a machine learning paradigm that assigns different weights to various types of misclassifications, guiding the model to minimize the "expected total misclassification cost" rather than simply the error rate (Elkan 2001). This approach has been shown to be effective in high-risk domains, such as medical diagnostics, where the cost of a false negative (missing a disease) is fatal, or financial fraud detection, where the loss from allowing a fraudulent transaction overwhelmingly exceeds the inconvenience of blocking a normal one (Dal Pozzolo et al. 2015). However, systematic studies formulating and applying this asymmetric CSL paradigm to the domain of multi-level information security and automated corporate document classification remain limited. To address this gap, we formulate information grading as a cost-sensitive classification problem.

Methodology

To control the asymmetric misclassification risks inherent in information grading, we proposed an automated document classification pipeline based on cost-sensitive learning. We designed a comprehensive classification pipeline consisting of five major stages: document preprocessing, security framework-based labeling, TF-IDF feature extraction, asymmetric cost matrix design, and conservative document-level aggregation.

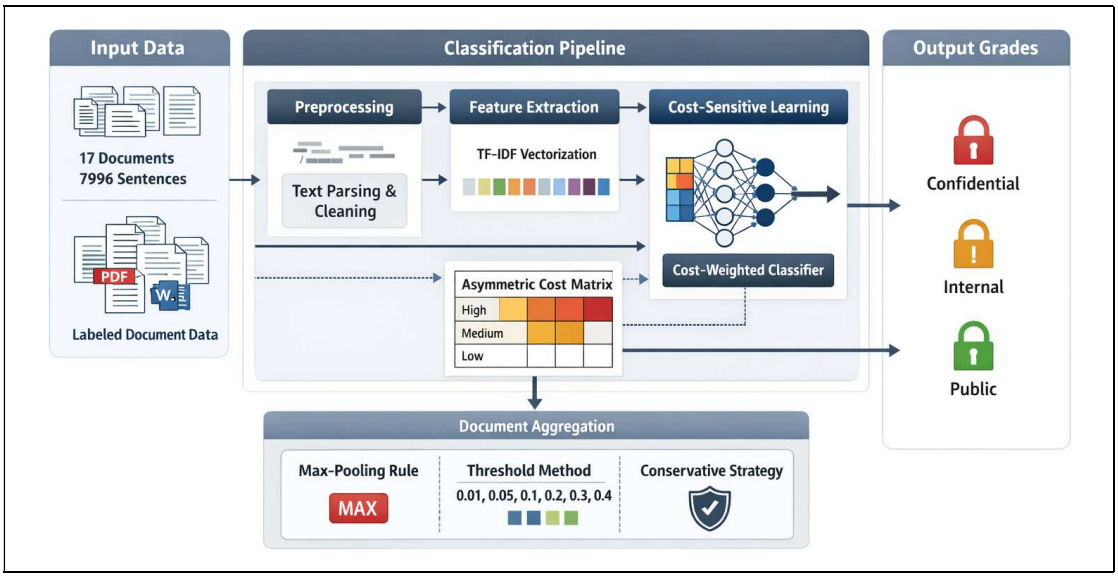


Figure 1. Overall process

Problem Definition

Standard multiclass classification models typically aim to maximize overall accuracy. However, in information security, risks vary substantially depending on the direction of misclassification. Downward misclassification leads to severe security breaches, whereas upward misclassification may cause operational inefficiencies. Therefore, we redefined information classification as a problem of minimizing expected cost rather than simply optimizing accuracy (Elkan 2001). Our primary goal was to algorithmically suppress downward misclassification of high-risk data.

Document Preprocessing and Sentence Segmentation

Processing an entire document as a single input may obscure localized security threats, as sections with different sensitivity levels can be diluted within the overall context. To address this, we used regular

expressions to segment documents into sentences, treating each sentence as the minimum unit of analysis. We also filtered out metadata noise and simple numerical sequences that did not provide meaningful contextual information for grading.

CSO Framework-based Classification Scheme

To ensure the model's practical applicability, we adopted a three-level information grading system based on the NIST FIPS 199 standard and data asset protection frameworks from a Chief Security Officer (CSO) perspective. This system was used to categorize data according to the potential adverse impact that an information breach could have on an organization. Following this convention, we assigned lower numerical values to grades with higher confidentiality. The definitions for each level are as follows:

- **Level 1 (Critical/High):** This level includes highly confidential information that, if disclosed, could cause serious damage to the organization's survival and competitiveness.
- **Level 2 (Medium):** This level includes internally sensitive information that could cause visible organizational damage if disclosed to unauthorized parties.
- **Level 3 (Low):** This level includes general information that would cause little or no harm to the organization even if publicly disclosed.

Feature Representation

To convert sentence text into numerical vectors, we applied the Term Frequency-Inverse Document Frequency (TF-IDF) technique. TF-IDF is a well-established baseline method in text classification studies, offering high computational efficiency in identifying domain-specific terminology characteristic of the security domain. This approach allowed us to evaluate the effectiveness of the cost-sensitive algorithm independently, without the confounding effects of complex deep embeddings.

Asymmetric Cost Matrix Design

The core of this methodology was the design of a cost matrix (C) that mathematically quantifies differences in misclassification costs across grades. We designed a cost function $C(i,j)$ that incorporates three distinct components: a distance-based penalty, an asymmetric penalty, and class importance weights. The element $C(i,j)$ represents the penalty incurred when a sentence with an actual grade of i is misclassified as j.

$$Cost\ Matrix\ C(i,j) = \alpha \cdot |i-j| + \beta \cdot I(i<j) \cdot e^{\gamma(j-i)} + \delta \cdot w_i$$

- **Distance-based penalty ($\boldsymbol{\alpha \cdot |i-j|}$):** A base penalty that increases linearly in proportion to the classification gap between the actual and predicted grades.
- **Asymmetric penalty ($\boldsymbol{\beta \cdot I(i<j) \cdot e^{\gamma(j-i)}}$):** This penalty activates exclusively in downward misclassification scenarios through the indicator function $I(i<j)$. As the grade gap $(j-i)$ widens, the penalty increases exponentially, discouraging the misclassification of highly confidential information into lower grades.
- **Class importance ($\boldsymbol{\delta \cdot w_i}$):** An intrinsic risk weight (w_i) adjusts the absolute base penalty upward when high-grade data are misclassified. (Here, α, β, γ, and δ are hyperparameters that control the weight of each respective component.)

The cost parameters $C(i,j)$ are defined to reflect the asymmetric risks inherent in information security grading. Unlike accuracy-driven models, this framework assigns higher penalties to downward misclassifications$(i<j)$ to prevent critical security breaches. Following ISO/IEC 27001 and NIST FIPS 199 standards, we applied a conservative weighting scheme where the cost of under-protecting Level 1 assets significantly outweighs the cost of over-protecting Level 3 data. This calibration ensures that model optimization aligns with the organization's risk appetite, prioritizing data breach prevention over administrative efficiency.

As shown in Figure 2, this matrix forced the model to reflect the disproportionate cost of treating confidential information as low-grade.

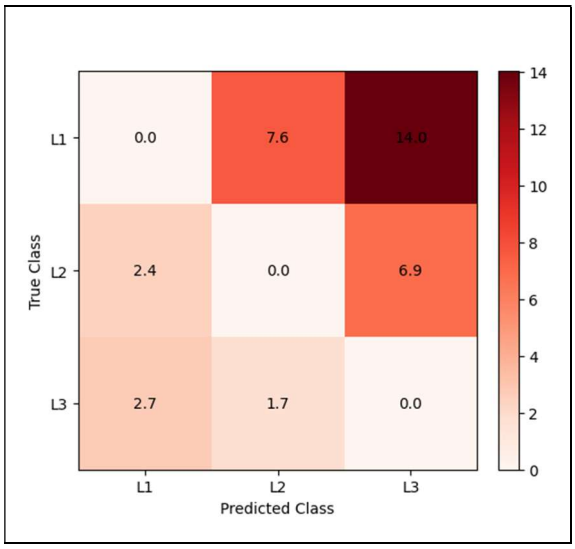


Figure 2. Cost matrix heatmap

Cost-sensitive Learning Strategy

To integrate the designed cost matrix into the training process, we adopted Logistic Regression as the base learner. This choice allowed us to evaluate the effect of cost-sensitive learning independently from model complexity. Logistic Regression is widely recognized as an effective model for TF-IDF-based text classification because of its robustness in high-dimensional sparse feature spaces (Manning et al. 2008). Its linear formulation provides stable and interpretable decision boundaries, making it suitable for controlled experimental settings.

To optimize the learner, we employed two strategies. First, we applied cost-informed sample weighting, in which training samples associated with higher misclassification costs were assigned larger weights, thereby increasing their influence during optimization. Second, we placed additional emphasis on high-grade document samples to ensure that the model effectively learned the characteristics of minority high-risk data.

Conservative Document-Level Aggregation

After sentence-level predictions were generated, we aggregated them to determine the final grade of the original document. Based on the standard assumption of Multi-Instance Learning (MIL), we adopted a Max-pooling aggregation rule. Even if a single confidential sentence is embedded within hundreds of general sentences, the entire document should be controlled at the highest required level. Therefore, the highest confidentiality grade among the sentence-level predictions was assigned as the final document grade.

Implementation

To verify the algorithmic validity and practical applicability of the proposed model, we conducted a two-phase evaluation of sentence-level performance. In the baseline model, the Logistic Regression based model, it reflected the importance of the document calculating according to the technical sensitivity system table set by the enterprise, and then the security level of the document finally determined according to the security regulations., we evaluated the performance and trade-offs of the classifier incorporating the cost matrix using an internally labeled sentence-level dataset (N = 1,833). Second, we performed a hold-out test using 120 raw corporate documents with established ground-truth labels to validate the effectiveness of the entire classification pipeline.

Sentence-Level Performance Analysis

To evaluate the effectiveness of the cost-sensitive approach, we compared our model against a standard Logistic Regression baseline. While the traditional accuracy-driven model achieved high overall performance, it consistently produced critical downward misclassifications due to the absence of a risk-weighting mechanism. By incorporating the asymmetric cost matrix, the proposed framework reduced the downward error rate by approximately 90%. These results indicate that a cost-aware approach provides a superior fail-safe mechanism for enterprise security, effectively mitigating high-impact data leakage risks at the expense of marginal overall accuracy. As illustrated in Figure 3, the baseline model achieved an overall accuracy of 84.51%, yet its downward error rate, defined as predicting high-risk data as a lower grade, reached 9.60% (a). Notably, the recall for Level 1 (critical) data was only 71.5%, revealing a vulnerability in security reliability (b).

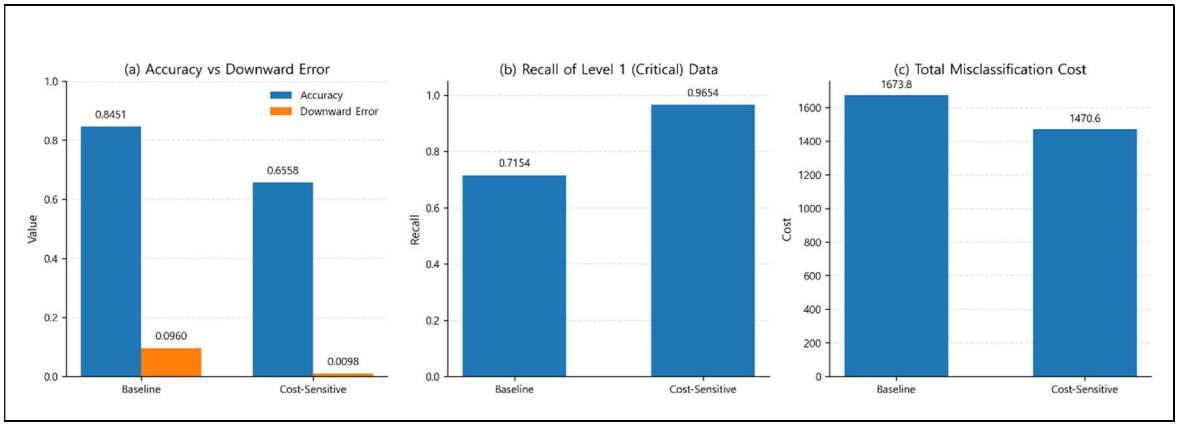


Figure 3. Overall performance: Baseline vs Cost-Sensitive

In contrast, our cost-sensitive model reduced the downward error rate to 0.98% and improved the recall of critical data to 96.54%, thereby effectively mitigating leakage risks (b). The total misclassification cost decreased from 1673.8 to 1470.6 (c). While the overall accuracy dropped to 65.58% due to the model's conservative prediction tendency, we interpreted this as a reasonable trade-off for improved control of data leakage risk.

Conservative Document-Level Aggregation

In the real-world document test using the majority-voting (MODE) aggregation method (Table 1), the baseline model achieved an accuracy of 48.33% but produced a downward error rate of 45.83%. Conversely, although the cost-sensitive model's accuracy dropped to 27.50%, it reduced the downward error rate to 0.03% and lowered the total cost by more than 48%, from 476.4 to 246.3.

Table 1. Performance Comparison (MODE)

	Baseline	Cost-Sensitive
Accuracy	0.4833	0.2750
Downward Error	0.4583	0.0333
Total Cost	476.4	246.3

Furthermore, we conducted threshold-adjustment experiments to verify the model's robustness (Figure 4). Even when the threshold varied from 1% to 30%, the proposed model maintained a stable decision boundary through parameter tuning. These results confirmed that the proposed framework provides consistent information protection without the need for manual post-hoc threshold adjustment.

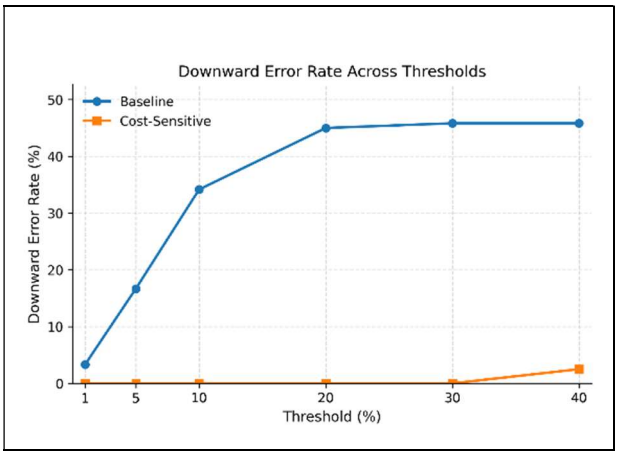


Figure 4. Downward Error Rate Across Thresholds

Conclusion

We proposed and validated an automated document classification pipeline that combined cost-sensitive learning with a conservative max-pooling aggregation rule to address the asymmetric misclassification risks in corporate information grading. The experimental results demonstrated that, despite a reduction in overall classification accuracy, the proposed model effectively suppressed the downward misclassification rate of high-risk confidential data to 0.00% and successfully reduced the total misclassification cost by over 54% compared to standard baselines.

The contributions of this study are twofold. First, from a theoretical perspective, we highlighted the structural limitations of traditional accuracy-driven information classification frameworks and redefined the task as a problem of minimizing expected misclassification costs. Specifically, by integrating a cost-weighted loss function with a Multi-Instance Learning (MIL)-based aggregation approach, we quantitatively showed that a cost-sensitive framework can provide robust security control even at the raw document level. Second, from a practical standpoint, the proposed model operationalizes the **Fail-Safe** principle demanded by Chief Security Officers (CSOs). It can serve as a reliable automated security classification system for cloud-based corporate information management environments that require large-scale document processing, as it fundamentally blocks critical data leakage without relying on manual threshold adjustment.

Despite these contributions, this study has several limitations. The model's conservative bias, which was intended to avoid critical security errors, resulted in over-protection, whereby general documents were classified into higher grades and some operational inefficiency was introduced. Additionally, we relied on TF-IDF-based feature extraction for computational efficiency.

To address these limitations, future research should focus on a Cost-Aware Ensemble framework that integrates deep pre-trained language models (e.g., BERT) with traditional feature extractors. By leveraging the complementary strengths of diverse classifiers through weighted ensemble voting, we aim to refine decision boundaries and mitigate over-protection while maintaining a strict security posture. This ensemble-driven approach will be further validated for its architectural scalability and tenant-specific policy adaptability within complex multi-tenancy cloud environments.

Furthermore, we plan to extend and verify the scalability of this cost-sensitive classification framework within complex distributed data architectures, such as multi-tenancy cloud environments.

Acknowledgements

This work was supported by the MSIT (Ministry of Science and ICT), Korea, under the Convergence security core talent training business support program(IITP-2025-RS-2023-00266605) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation)

References

Dal Pozzolo, A., Caelen, O., Johnson, R. A., and Bontempi, G. 2015. "Calibrating Probability with Undersampling for Unbalanced Classification," in *Proceedings of the IEEE Symposium Series on Computational Intelligence*, pp. 159-166.

Elkan, C. 2001. "The Foundations of Cost-Sensitive Learning," in *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pp. 973-978.

Evans, A., and Reeder, F. 2004. "Standards for Security Categorization of Federal Information and Information Systems," FIPS PUB 199, National Institute of Standards and Technology.

Ignaczak, L., et al. 2026. "Establishing a Value and Sensitivity-Based Information Classification Framework using NLP Models," *Journal of Information Security and Applications* (10:2), pp. 117-130.

International Organization for Standardization (ISO). 2022. "Information security, cybersecurity and privacy protection — Information security management systems — Requirements," ISO/IEC 27001:2022.

ISMS. online. (2025). ISO/IEC 27001:2022 Annex A 5.12 – Classification of Information.

National Institute of Standards and Technology (NIST). 2023. "Cybersecurity Resource and Reference Guide," NIST IR 8496.

Rāts, J. 2021. "Tunable Document Classification Framework for Enterprise Content Management Systems," in *Business Information Systems Workshops*, Springer, pp. 50-61.

Tan, L., et al. 2024. "Content-Based Document Classification for Sensitive Information Identification in Enterprise Repositories," *Applied Sciences* (14:4), p. 1565.

Tari, Z. 2014. "Security and Privacy in Cloud Computing," *IEEE Cloud Computing* (1:1), pp. 54-57.

Toulas, B. 2024. "National Public Data Confirms Massive Data Breach Impacting Millions," *BleepingComputer*. [Online]. Available: https://www.bleepingcomputer.com/news/security/national-public-data-confirms-massive-data-breach-impacting-millions/

Allam, H., Makubvure, L., Gyamfi, B., Nyarko Graham, K., and Akinwolere, K. 2025. "Text Classification: How Machine Learning Is Revolutionizing Text Categorization," Information (16:2), Article 130.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval.* Cambridge University Press.

C2

How AIGC Disclosure Labels Drive Consumers' UGC Compensatory Search Behavior

Zhenhan Qin, Xusen Cheng, Xinhe Qu, Ang Zeng, Yijun Yan (Renmin University of China).

AIGC has drastically lowered the aesthetic threshold for e-commerce visual content, making AI-generated product images an indispensable strategy for merchants toaccelerate new product launches. However, government-mandated AIGC disclosure policies – requiring explicit labeling of AI-generated content – trigger consumers' perception of "authenticity violation." Grounded in the Compensatory Behavior Theory, this study investigates consumers' decision-making pathways following the devaluation of official information sources. Using a 2 × 2 between-subjects experimental design, we deployed a self-developed simulated e-commerce platform and JavaScript-based tracking to capture objective behavioral data. Preliminary findings reveal that AIGC disclosure labels induce a shift in search focus toward User-Generated Content (UGC), thereby reshaping the weight structure of consumers' decision-making.

KEY WORDS

Generative AI, AI Disclosure Labels, Compensatory Search, User-Generated Content (UGC), Weight Reconfiguration

How AIGC Disclosure Labels Drive Consumers' UGC Compensatory Search Behavior

Qin, Zhenhan[a], Cheng, Xusen[b], Qu, Xinhe[c], Zeng, Ang[d] and Yan, Yijun[e]
[a] School of Information, Renmin University of China
E-mail: 3030916181@qq.com
[b] School of Information, Renmin University of China
mail: xusen.cheng@ruc.edu.cn
[c] School of Information, Renmin University of China
E-mail: quxinhe2020@ruc.edu.cn
[d] School of Information, Renmin University of China
E-mail: zengang@ruc.edu.cn
[e] School of Information, Renmin University of China
E-mail: yanyijun@ruc.edu.cn

Abstract

AIGC has drastically lowered the aesthetic threshold for e-commerce visual content, making AI-generated product images an indispensable strategy for merchants to accelerate new product launches. However, government-mandated AIGC disclosure policies—requiring explicit labeling of AI-generated content—trigger consumers' perception of "authenticity violation." Grounded in the Compensatory Behavior Theory, this study investigates consumers' decision-making pathways following the devaluation of official information sources. Using a 2 × 2 between-subjects experimental design, we deployed a self-developed simulated e-commerce platform and JavaScript-based tracking to capture objective behavioral data. Preliminary findings reveal that AIGC disclosure labels induce a shift in search focus toward User-Generated Content (UGC), thereby reshaping the weight structure of consumers' decision-making.

Keywords: Generative AI; AI Disclosure Labels; Compensatory Search; User-Generated Content (UGC); Weight Reconfiguration

Problem Background

The rapid iteration of generative AI technology has significantly reduced barriers to creating e-commerce visual content, with AI-generated product ads emerging as an inevitable trend for merchants to enhance new product launch efficiency and diversify visual presentation (Bawack et al., 2022). However, government mandates requiring e-commerce platforms to affix exclusive labels to AI-generated product ads have disrupted consumers' prior expectations of the authenticity of official advertising content, eroding trust in both the ads and the organizations behind them (Koning et al., 2025).

Existing literature primarily focuses on "algorithm aversion" and trust erosion caused by AI disclosure. In contrast, this study posits that when consumers perceive authenticity deficits in official ads, they do not passively abandon purchase intentions but instead engage in compensatory search behavior — strengthening verification of third-party alternative cues (e.g., UGC) to mitigate perceived risk. This "attention reallocation" is expected to fundamentally reshape the weight structure of e-commerce evaluation systems (Schmitt et al., 2024).

Theoretical Framework & Hypotheses

Compensatory Behavior Triggered by AI Advertising Disclosure

Disclosing the AI-generated nature of ads elicits consumers' perceptions of "authenticity violation" and "dehumanization," which significantly reduce brand trust and purchase intent (Bui, 2025; Belanche et al., 2025). Consumer decision-making inherently involves integrating the weights of multiple information sources. According to compensation strategy (Payne, 1993) , when AIGC disclosure devalues the diagnostic value of official ads, consumers engage in compensatory information search to reduce decision risk, actively seeking more credible alternative sources.

UGC—real experiential content created by ordinary consumers in e-commerce contexts—exhibits high authenticity, strong emotional resonance, and rich informational dimensions, making it the primary alternative information source for consumers (Li et al., 2020). We thus propose:

H1: AIGC disclosure labels significantly increase consumers' search effort toward UGC.

Moderating Role of UGC Information Quality

UGC's informational value is heterogeneous, and its quality directly determines the effectiveness of this alternative source. According to the Elaboration Likelihood Model (ELM), when consumers are in a high-motivation search state, "argument quality" becomes the core moderator of attitude change (Bhattacherjee & Sanford, 2006). High-quality UGC serves as a reliable carrier of authentic product signals, while low-quality UGC constitutes invalid or even negative signals. Thus:

H2: The positive restorative effect of UGC search effort on purchase intent is stronger when UGC quality (W) is higher.

H3: In contexts with AIGC disclosure labels, the effect size of UGC quality on decision-making significantly exceeds that of official ad quality.

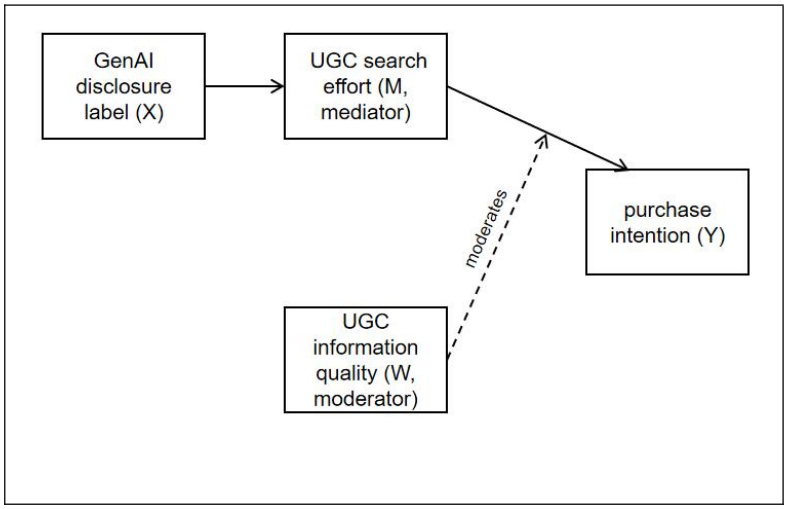


Figure 1. Research Model

Research Methodology

Mathematical Models

We test hypotheses using a series of regression models:

To test H1, we estimate UGC search effort via regression model 1 incorporating the AIGC disclosure label as the independent variable.

$$M_i = \alpha_0 + \alpha_1 X_i + \epsilon_{1i} (1)$$

To verify H2 and H3, we employ grouped regression and interaction models to examine the moderating effect of UGC quality and compare the relative effect sizes of UGC quality and official ad quality using model 2.

$$Y_i = \gamma_0 + \gamma_1 X_i + \gamma_2 M_i + \gamma_3 W_i + \gamma_4 (X_i \times W_i) + \epsilon_{2i} (2)$$

Experimental Design

A 2 (AIGC Disclosure Label: Present vs. Absent) × 2 (UGC Quality: High vs. Low) between-subjects design was adopted. Participants (N = 400) were randomly assigned to four experimental groups (n = 100 per group). A simulated e-commerce webpage was developed, and JavaScript-based tracking was embedded to capture the relative dwell time ratio (time spent in the comment section vs. main image section), obtaining a search weight index $M_i = T_{ratio} = \frac{T_{ugc}}{T_{ad}+T_{ugc}}$ as an objective metric of search effort. Post-purchase, purchase intent was measured using a validated scale adapted from prior literature.

Potential Contributions

Existing research on AI-generated advertising has predominantly centered on exploring its negative effects on consumer trust, perceived quality, and attitudes. However, there is a paucity of discussions regarding how consumers respond and adjust their information-seeking strategies in the context of declining AI credibility. The findings of this study extend the post-effect research on AI advertising effectiveness, enrich the theoretical accounts of information-seeking behavior in AI-enabled environments, and offer a novel perspective for understanding the consumer decision-making mechanism within the human-machine collaborative content ecosystem.

Acknowledgements

Xusen Cheng is the corresponding author. We would like to thank Key Projects of Philosophy and Social SciencesResearch of Chinese Ministry of Education (Grant No.23JZD022) and the National Natural ScienceFoundation of China (Grant No. 72434006, 72271236) for providing funding for part of thisresearch.

References

Bawack, R., Wamba, S., Carillo, K., and Akter, S. 2022. "Artificial Intelligence in E-Commerce: A Bibliometric Study and Literature Review," Electronic Markets (32:1), pp. 297-338.

Belanche, D., Ibáñez-Sánchez, S., Jordán, P., and Matas, S. 2025. "Customer Reactions to Generative AI vs. Real Images in High-Involvement and Hedonic Services," International Journal of Information Management (85), 102954.

Bui, H. 2025. "Examining the Effect of AIGC Advertising Involvement Disclosure on Advertising Value and Purchase Intentions," Journal of Research in Interactive Marketing, [Online First], https://doi.org/10.1108/jrim-02-2025-0066.

Koning, B., and Voorveld, H. 2025. "Disclaimer! This Content Is AI-Generated: How AIGC-Disclosures Influence Trust in Advertisements and Organizations," Journal of Interactive Advertising (25:2), pp. 240-253.

Li, S., Zhang, Y., Yu, Z., and Liu, F. 2020. "Economical User-Generated Content (UGC) Marketing for Online Stores Based on a Fine-Grained Joint Model of the Consumer Purchase Decision Process," Electronic Commerce Research, pp. 1-30.

Bhattacherjee, A., and Sanford, C. 2006. "Influence Processes in Technology Acceptance," MIS Quarterly (30:4), pp. 805-825.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. Cambridge University Press.

C2

Analysis of the Influences of AI on the Media Industry with Story Telling

Hui Liu, Jinhwa Kim (Sogang University)

This study investigates how AI influences the media industry by analyzing interactions between AI and media. Data was collected from Naver News, Twitter, Naver Blogs, and Instagram, covering both news and social media. Using a storytelling approach, the study applies text mining techniques such as dynamic topic modelling, network analysis, and association rule mining. Results highlight the mediating role of digital and platform technologies in the AI–media relationship. Moreover, findings suggest that emerging technologies like the metaverse and data-driven systems, enhanced by AI, will significantly shape the future of the media industry.

KEY WORDS

AI, Media, Text Mining, Network Analysis, Association Rules Analysis

Tracing AI's Influence on the Media Industry: A Storytelling-Based Text Mining Approach

Liu Hui[a] and Kim Jinhwa[b]
[a] Graduate School of AI·SW, Sogang University
E-mail: yoohye@sogang.ac.kr
[b] School of Business, Sogang University
E-mail: jinhwakim@sogang.ac.kr

Abstract

This study examines how artificial intelligence (AI) influences the media industry by analyzing large-scale data related to AI and media. Using textual data collected from Naver News, Twitter, Naver Blogs, and Instagram, the study adopts a storytelling-based text mining approach that integrates frequency analysis, TF-IDF, dynamic topic modeling, network analysis, and association rule mining. The analysis traces shared vocabularies, evolving topics, structural relationships, and directional pathways linking AI and media. The results show that AI–media is centered on technology, digital transformation, enterprise activity, and platform-based services. More importantly, AI influences the media industry not through a simple direct effect, but through digital, platform, enterprise, and data-driven pathways. These findings suggest that AI is reshaping the media industry as an evolving socio-technical system. The study contributes by offering an integrated explanation of the structural mechanisms linking AI and media and by demonstrating the usefulness of a storytelling-based text mining framework for understanding industry transformation.

Keywords: Text Mining, Network Analysis, Association Rules Analysis, Story Telling

Introduction

Artificial intelligence (AI), together with advances in deep learning, machine learning, and big data, is increasingly reshaping the media industry. In particular, AI has improved media services by enabling personalized content recommendation, enhancing decision-making, increasing operational efficiency, and strengthening audience engagement (Anagnostou et al. 2022; Shields 2018; Chan-Olmsted 2019; Duan et al. 2019). As a result, media has increasingly shifted from traditional one-way delivery to more interactive and personalized forms of content production, distribution, and consumption.

This transformation is especially visible on digital media platforms such as Netflix and YouTube, where AI is used to analyze user preferences and provide tailored content. Prior studies have examined AI's role in journalism, content delivery, entertainment, and recommendation systems, while also emphasizing concerns related to bias, misinformation, privacy, and ethics (Budzinski et al. 2021; Broussard et al. 2019; de-Lima-Santos and Ceron 2021; Trattner et al. 2022). Although these studies provide important insights into the applications and risks of AI in media, they have paid relatively less attention to how AI structurally influences the media industry through evolving digital, platform-based, enterprise, and data-related mechanisms across large-scale textual discourse.

To address this gap, this study examines how AI is connected to the media industry by identifying major themes, structural relationships, and influence pathways in AI- and media-related discourse. In this study, storytelling refers to a data-driven analytical narrative that reconstructs the relationship between AI and media by tracing shared vocabularies, evolving topics, network structures, and association

pathways. Specifically, this study asks: What themes, network structures, and association pathways characterize the relationship between AI and the media industry? To answer this question, the study applies multiple text-mining methods, including frequency analysis, TF-IDF, dynamic topic analysis, network analysis, and association rule analysis.

This study makes two contributions. First, it provides an integrated explanation of how AI influences the media industry through technology, platforms, digital transformation, enterprise activity, and data-driven services. Second, it shows that a storytelling-based combination of text-mining methods can reveal not only major themes, but also the structural and relational mechanisms linking AI and media. In doing so, the study moves beyond a descriptive account of AI applications in media and offers a more systematic understanding of how AI is reshaping the media industry.

Conceptual Background

AI Applications in the Media Industry

The media industry includes traditional media, such as newspapers, television, and radio, as well as digital media, including social media, blogs, and streaming platforms. Unlike traditional one-way communication, digital media enables interactivity and user-generated content, and AI has become a major driver of this transformation.

Previous studies show that AI enhances the media industry through personalized recommendation, audience engagement, content creation, content management, and workflow automation (Chan-Olmsted 2019). For example, platforms such as Netflix and YouTube use AI to provide personalized content recommendations (Wayne 2022), while AI is also applied in news production through fact-checking and data analysis (Nakov et al. 2021). In addition, AI is increasingly used in personalized advertising and other commercial strategies.From an operational perspective, AI improves both efficiency and user experience by better matching content with user preferences (Gentzkow 2018). More broadly, AI is transforming how media content is produced, distributed, and consumed, making it an important technological force in the ongoing digital transformation of the media industry.

Research Methodologies

Data Collection and Preprocessing

This study collected AI- and media-related textual data over a one-year period from multiple online platforms, including Naver News, Twitter, Naver Blogs, and Instagram. Using the keywords “AI” and “media,” a total of 83,752 AI-related documents and 48,254 media-related documents were gathered. After collection, the text data were preprocessed by removing stop words and irrelevant terms to improve the reliability of the analysis. The cleaned words were then used for frequency analysis, TF-IDF, dynamic topic analysis, network analysis, and association rule analysis. These methods were organized as a sequential storytelling framework: frequency and TF-IDF analyses first identified shared vocabularies between AI and media, dynamic topic analysis traced how major themes evolved over time, network analysis revealed the structural relationships among key concepts, and association rule analysis identified the directional pathways through which AI is linked to media.

Frequency and TF-IDF Analysis

Frequency analysis was used to identify the most commonly occurring keywords in the AI and media corpora and to examine overlapping terms between the two domains. To complement this, TF-IDF analysis was applied to identify terms that were not only frequent but also relatively distinctive within each corpus. Together, these two methods provide a basic understanding of the major keywords and lexical characteristics shaping AI- and media-related discourse, thereby forming the first stage of the storytelling process.

Dynamic Topic Analysis

Dynamic topic analysis was employed to identify major themes in the collected text data and to examine how these themes changed over time. The optimal number of topics was determined using coherence and perplexity, which are widely used evaluation measures in topic modeling. After selecting the final number of topics, topic trends were analyzed longitudinally to capture shifts in emphasis within AI–media discourse. In addition, cosine similarity was used to examine the relationships among the identified topics. This stage captures the temporal evolution of AI–media discourse.

Network Analysis

Network analysis was conducted to examine the structural relationships among keywords in the AI–media discourse. In this study, words were treated as nodes and their co-occurrences as links. To identify influential and bridging terms, three centrality measures were considered: degree centrality, eigenvector centrality, and betweenness centrality. Degree centrality indicates the number of direct connections a node has, eigenvector centrality reflects the importance of connected nodes, and betweenness centrality captures the extent to which a node functions as a bridge within the network. Community analysis was also performed to identify major clusters of closely related terms. This stage reveals the structural organization of the relationship between AI and Media.

Association Rule Analysis

Association rule analysis was used to identify the main pathways through which AI is linked to the media industry. This method helps reveal sequential relationships among keywords by estimating the confidence of each path. In this study, total confidence was calculated by multiplying the confidence values across consecutive transitions in each pathway. Through this approach, the study identifies both direct and indirect connections between AI and media-related concepts. This final stage of the storytelling framework highlights the directional pathways linking AI and media.

Research Processes and Results from the Analysis

Research Processes

Figure 1 presents the overall research process, which also serves as the storytelling framework of this study. The study first collected and preprocessed the data by removing stop words and irrelevant terms. It then applied frequency analysis and TF-IDF to identify key terms related to AI and media, followed by dynamic topic analysis to extract major themes and examine their changes over time. Next, network analysis was conducted to explore the structural relationships among keywords and visualize them as an interconnected discourse network. Finally, association rule analysis was performed to identify the directional pathways through which AI is linked to the media industry.

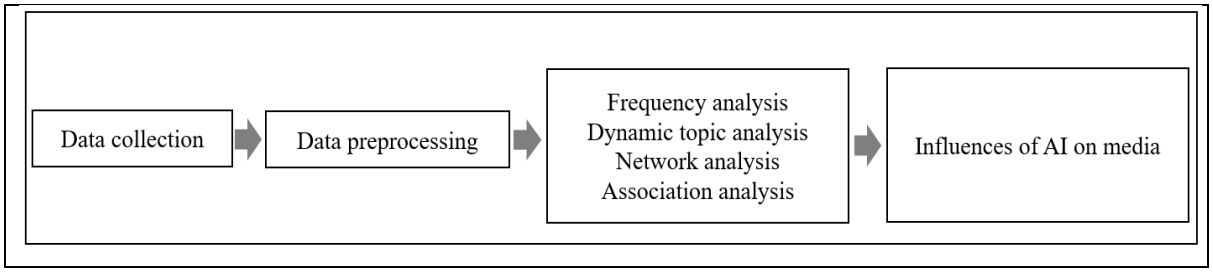


Figure 1. Research processes

As of August 24, 2023, a total of 48,254 media-related documents and 83,752 AI-related documents were collected over one year from various sources, including Naver News, Twitter, Naver Blogs, and Instagram, using the keywords “AI” and “media.” The collected text data were preprocessed and analyzed using frequency analysis, TF-IDF, dynamic topic analysis, network analysis, and association rule analysis.

Results from Frequency and TF-IDF Analyses

To begin the storytelling process, frequency and TF-IDF analyses were used to identify the shared lexical foundations of AI–media discourse. The analyses identified several shared keywords across the AI and media corpora, including "Technology," "AI," "Enterprise," "Business," and "Global," reflecting the convergence of the two fields around technological and industrial transformation. However, AI was more closely associated with digital services and platforms, whereas media emphasized content, culture, and broadcasting. These results suggest that AI–media discourse is driven by technology and enterprise change, while media-specific discussions remain more content-oriented.

Results from Dynamic Topic Analysis

Building on the shared keywords identified above, dynamic topic analysis was conducted to trace how AI–media discourse evolved over time. To identify the optimal number of topics, coherence and perplexity were used, and 11 topics were selected. The dynamic topic analysis revealed that Topic 2 (Industrial Innovation), Topic 4 (Education and Digital Innovation), and Topic 6 (Automotive Industry) increased toward the end of the observation period, indicating growing attention to industrial transformation, education, and the automotive sector. In addition, Topic 3 (Cultural Events) and Topic 9 (AI and Digital Media) became more prominent after May 2022, suggesting rising interest in AI-related changes in cultural and media activities. In particular, Topic 4 reflects the growing convergence between AI, digital innovation, and education, as illustrated in Figure 2.

The cosine similarity analysis further showed meaningful relationships among topics. Topics 3 and 9 were relatively close, implying a strong connection between cultural content and AI-driven digital media. Likewise, Topics 2, 7, and 11 were closely related, suggesting that AI is strongly associated with industrial innovation, enterprise investment, and digital technology services. Overall, the results indicate that AI-related media discourse increasingly centers on industrial innovation, digital services, education, and cultural-media convergence.

Table 1. Results of dynamic topic analysis

Topic	Topic Label	Representative Keywords
1	Art and Social Platform	art, social media, brand, transaction, YouTube, big data
2	Industrial Innovation	industry, technology, digital, enterprise, innovation, AI
3	Cultural Events	culture, media, event, exhibition, tourism, contents
4	Education and Digital Innovation	education, program, university, student, digital, start-up
5	Global Pandemic	COVID-19, global, media, crisis, China, economic
6	Automotive Industry	automobile, mobility, electric vehicle, robot, autonomy, safety
7	Enterprise Investment	investment, start-up, enterprise, market, platform, venture
8	Democratic Governance	election, citizen, government, media, policy, democracy
9	AI and Digital Media	contents, media, broadcast, video, AI, metaverse
10	Regional Business Development	region, business, revenue, profit, operation, promotion
11	Digital Technology Service	platform, customer, digital, AI, data, cloud

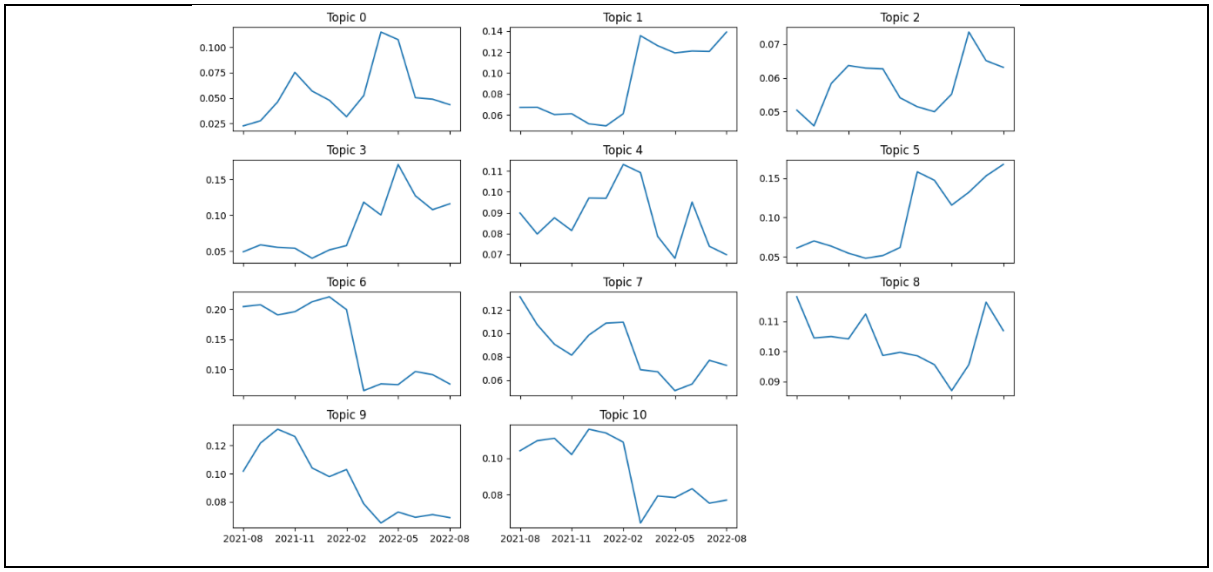


Figure 2. Dynamic topics changes over time

Network Analysis

To further reveal the structural organization of AI–media discourse, network analysis was conducted using degree centrality and betweenness centrality. Across the centrality measures, Technology emerged as the most central node, followed by AI, Global, Enterprise, Industry, Business, and Platform. These results indicate that technology-related concepts occupy a central position in the AI–media network. In addition, terms such as Metaverse, Reality, Contents, Service, and Digital also showed high centrality, suggesting the growing importance of immersive technology, digital transformation, and content-based services in the media industry. Figure 3 visualizes the network structure based on the Top50 degree centrality and betweenness centrality results.

Community analysis further identified four major clusters in the network. The first cluster relates to the business ecosystem, including enterprise, investment, and global market expansion. The second cluster represents technical and structural aspects, such as AI, data, system integration, and solutions. The third cluster reflects societal and economic development, covering strategy, policy, government, education, and COVID-19-related issues. The fourth cluster centers on digital and cultural experiences, including media, contents, virtual platforms, metaverse, and user experience. Overall, the network results suggest that AI–media discourse is embedded in a broader socio-technical ecosystem structured around business, technology, policy, and digital cultural experience.

Table2. Community analysis of degree centrality and betweenness

Community	Theme	Representative Terms
1	Business Ecosystem	enterprise, startup, investment, company, global market
2	Technical Infrastructure	AI, data, analysis, system, integration, solution
3	Societal and Policy Development	strategy, government, policy, education, economy, COVID-19

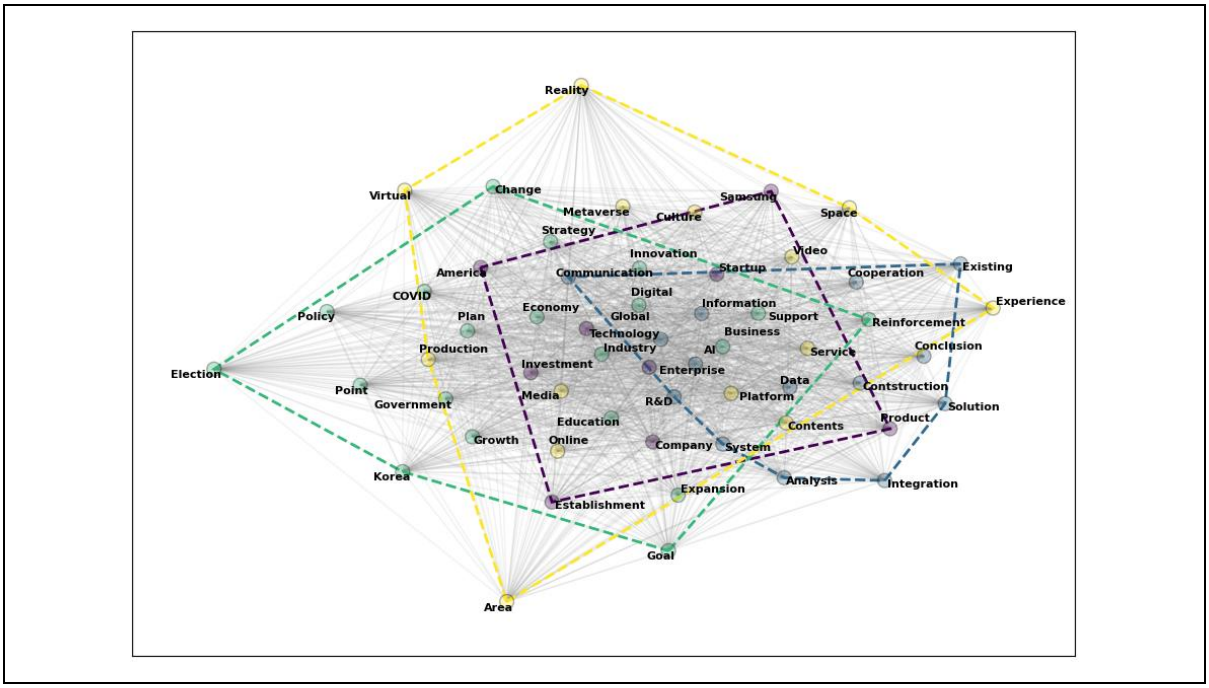


Figure 3. Visualization based on Top50 DC and BC

Association Rule Analysis Results

Finally, association rule analysis was conducted to identify the directional pathways through which AI influences the media industry. Rather than focusing solely on co-occurrence relationships, this analysis reveals how AI is connected to media through sequential and mediated linkages among key concepts.

The results identify several representative pathways. Direct pathways such as AI → Platform → Media, AI → Global → Media, and AI → Digital → Media indicate that AI is closely associated with platform-based services, globalization, and digital transformation in the media industry. In addition, more complex pathways were identified, including AI → Technology → Digital → Media, AI → Enterprise → Platform → Media, and AI → Data → Platform → Media, suggesting that AI operates through multiple intermediate layers.

These pathways can be broadly grouped into three types. First, platform-mediated pathways (e.g., AI → Platform → Media; AI → Data → Platform → Media) indicate that AI influences media primarily through platform-based ecosystems, where data and user interaction play a central role. Second, technology- and digital-mediated pathways (e.g., AI → Technology → Digital → Media; AI → Digital → Media) suggest that AI reshapes the media industry through digital infrastructure and technological advancement. Third, socio-economic pathways (e.g., AI → Enterprise → Platform → Media; AI → Global → Media) highlight the role of enterprise activity, investment, and globalization in linking AI to media transformation.

Overall, these results suggest that AI influences the media industry not through a simple direct effect, but through a set of interconnected and mediated pathways centered on platforms, digital technologies, enterprise systems, and data-driven services. This finding highlights the importance of understanding AI as part of a broader socio-technical ecosystem rather than as an isolated technological factor.

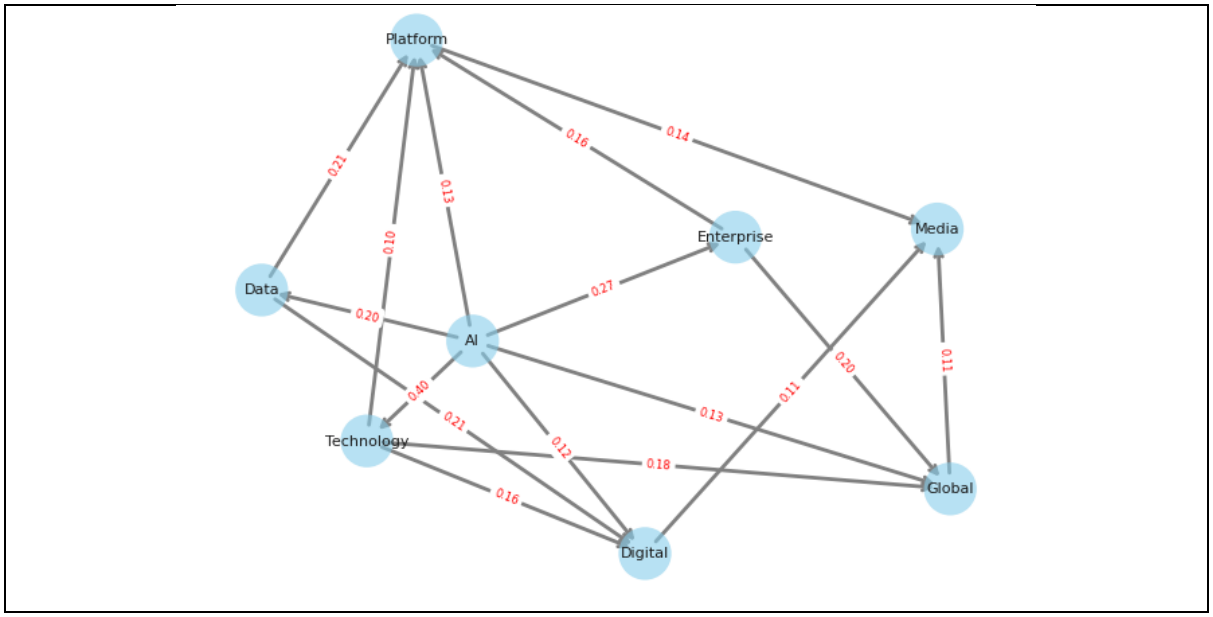


Figure 4. Top 10 Confidence Paths from AI to Media

Conclusion

This study investigates how artificial intelligence (AI) influences the media industry by analyzing large-scale AI–media-related discourse using a storytelling-based text mining approach. The findings show that the relationship between AI and the media industry is characterized by interconnected themes, structured networks, and, most importantly, mediated pathways rather than simple direct effects.

The results indicate that AI–media discourse is centered on technology, digital transformation, enterprise activity, and platform-based services. Dynamic topic analysis reveals an increasing focus on industrial innovation, digital services, and cultural-media convergence over time. Network analysis further demonstrates that AI–media discourse is organized around a socio-technical ecosystem composed of business, technology, policy, and digital cultural experience. Most importantly, association rule analysis shows that AI influences the media industry through a set of interconnected pathways, including platform-based, digital, technological, enterprise, and data-driven mechanisms, rather than through a direct one-to-one relationship.

The theoretical contribution of this study lies in showing that the relationship between AI and the media industry should be understood as a mediated and evolving socio-technical process. Rather than viewing AI as a standalone technological factor, this study demonstrates that AI reshapes the media industry through multiple intermediate layers, including platforms, digital infrastructures, enterprise systems, and data-driven services. In this sense, the study moves beyond descriptive analysis and provides an integrated explanation of the structural mechanisms linking AI and media.

Practically, the findings offer several implications. For media firms, the results suggest that AI investment should extend beyond content generation to include platform capabilities, data integration, and digital infrastructure development. For content producers and broadcasters, the growing importance of AI-assisted personalization, multi-platform distribution, and audience analytics is highlighted. For policymakers and educators, the findings underscore the need to strengthen digital capability, platform governance, and data literacy in response to the AI-driven transformation of the media industry.

This study has several limitations. First, the analysis is based on Korean-language data from specific online platforms, which may limit the generalizability of the findings. Second, the study focuses on textual discourse and does not incorporate other forms of media data such as images or videos. Future research could extend this study by incorporating multi-modal data and cross-national comparisons to provide a more comprehensive understanding of AI’s impact on the media industry.

References

Anagnostou, M., Karvounidou, O., Katritzidaki, C., Kechagia, C., Melidou, K., Mpeza, E., Konstantinidis, I., Kapantai, E., Berberidis, C., and Magnisalis, I. 2022. "Characteristics and Challenges in the Industries Towards Responsible Ai: A Systematic Literature Review," Ethics and Information Technology (24:3), p. 37.

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., and Chuan, C.-H. 2019. "Artificial Intelligence and Journalism," Journalism & mass communication quarterly (96:3), pp. 673–695.

Budzinski, O., Gaenssle, S., and Lindstädt-Dreusicke, N. 2021. "The Battle of Youtube, Tv and Netflix: An Empirical Analysis of Competition in Audiovisual Media Markets," SN Business & Economics (1:9), p. 116.

Chan-Olmsted, S. M. 2019. "A Review of Artificial Intelligence Adoptions in the Media Industry," International journal on media management (21:3-4), pp. 193–215.

de-Lima-Santos, M.-F., and Ceron, W. 2021. "Artificial Intelligence in News Media: Current Perceptions and Future Outlook," Journalism and media (3:1), pp. 13–26.

Duan, Y., Edwards, J. S., and Dwivedi, Y. K. 2019. "Artificial Intelligence for Decision Making in the Era of Big Data–Evolution, Challenges and Research Agenda," International journal of information management (48), pp. 63–71.

Gentzkow, M. 2018. "Media and Artificial Intelligence," Toulouse Network for Information Technology Whitepaper).

Nakov, P., Corney, D., Hasanain, M., Alam, F., Elsayed, T., Barrón-Cedeño, A., Papotti, P., Shaar, S., and Martino, G. D. S. 2021. "Automated Fact-Checking for Assisting Human Fact-Checkers," arXiv preprint arXiv:2103.07769).

Shields, R. 2018. "What the Media Industry Really Thinks About the Impact of Ai."

Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., Opdahl, A. L., Tessem, B., Borch, N., and Fjeld, M. 2022. "Responsible Media Technology and Ai: Challenges and Research Directions," AI and Ethics (2:4), pp. 585–594.

Wayne, M. L. 2022. "Netflix Audience Data, Streaming Industry Discourse, and the Emerging Realities of 'Popular'television," Media, Culture & Society (44:2), pp. 193–209.

C2

Intrusive or Effective? The Trade-Off between Brand Recall and Viewer Satisfaction of In-Game Advertising in eSports

Sze Ming Chow, Yan Sun, Sung-Byung Yang (Kyung Hee University)

As eSports evolves into a globally institutionalized ecosystem, this study investigates how two In-Game Advertising (IGA) formats (overlay and embedded ads) influence perceived ad intrusiveness across varying levels of game intensity. Grounded in the LC4MP framework, the research examines the cognitive conflict between goal-directed eSports viewing and the processing of advertisements. A 2x2 between-subjects experiment is conducted to examine the impact of IGA formats and game intensity on perceived ad intrusiveness and how these factors shape the trade-off between brand recall and viewer satisfaction. The study aims to provide strategic insights for platforms to maximize monetization through effective ad placement while maintaining a positive audience experience.

KEY WORDS

In-Game Advertising, Perceived Ad Intrusiveness, eSports, LC4MP, Brand Recall

KITS·ICEC·KIISS
2026 통합 국제학술대회

C3 Digital Finance & Platform Economy

Session Chair : Kyung Young Lee (Dalhousie Univ.)



C3

Understanding the Endorsement Effect of Digital Virtual Human on Investment Intention Towards Crowdfunding-Based on Elaboration Likelihood Model

Jiajie Chu (Southwestern University of Finance and Economics)

The application of AI has surged in crowdsourcing and crowdfunding platforms, driving not only business but also charitable and philanthropic efforts. Despite the diverse AI applications available, these platforms often struggle to fully understand their utility.The primary difficulty lies in integrating AI tools on donation-based platforms with low profitability, where the adoption process remains underexplored. To explore factors affecting users' intentions, this study applies Elaboration likelihood model to explain how digital virtual human affect user's intestment intention in the context of donation crowdfunding . According to existing online trust model, the factors in this study are categorized into two dimensions: user-based and platform-based. This study suggests self-reference, para-social interaction, perceived congruence , transaction convenience, empathy and perceived credibility provide a strong explanation for investment intention in donation crowdfunding. The results of this research provided a theoretical foundation for future academic studies as well as practical guidance for rapid development of digital virtual human.

KEY WORDS

Digital Virtual Human, Donation-based Crowdfunding, Elaboration Likelihood Model, Empathy, Perceived Credibility

C3

Probabilistic Earnings Forecasting Using Large Language Models in Digital Platforms

Xiaolin Gan, Xincan Qiu, Chenglong Zhang (College of Business Administration, Kookmin University)

This study examines large language models (LLMs) as core components of AI-driven prediction systems in digital platforms. Unlike traditional models that rely primarily on structured numerical inputs, LLMs are capable of transforming financial information into predictive signals through semantic understanding and contextual reasoning. We develop a probabilistic forecasting framework to evaluate the performance of LLMs in earnings direction prediction. The prediction task is formulated as a binary classification problem based on structured financial statement data. Model performance is assessed using a multi-dimensional metric system, including LogLoss, Brier Score, AUC, and calibration measures such as Expected Calibration Error (ECE). Using out-of-sample evaluation, we find that LLM-based predictions exhibit statistically meaningful predictive power and stable probabilistic calibration. Compared to benchmark models, including logistic regression, LLMs demonstrate improved performance in both ranking ability and probability quality. These findings contribute to the literature by highlighting the role of LLMs as probabilistic prediction modules embedded in digital platform decision systems and by emphasizing the importance of calibration in AI-driven forecasting.

KEY WORDS Digital Platforms, LLM, Decision Systems, Earnings Prediction, Probabilistic Forecasting

C3

Analysis of Solana Validator Staking Amount: A Machine Learning Approach

Ke Li, Chenglong Zhang (College of Business Administration, Kookmin University)

This study adopts a data-driven approach, integrating traditional regression and machine learning methods to investigate the determinants of validator staking amounts in the Solana blockchain. Using a dataset of 779 active validators, explanatory variables are categorized into economic, brand-related, service quality, and staking structure factors. The analysis compares five machine learning regression models to assess their performance and analyze the factor influencing validator staking in Solana. Model performance is evaluated using MAE, RMSE, and R^2, while interpretability is examined through Feature Importance and SHAP. The results show that TabPFN outperforms other individual models and traditional regression in predictive accuracy and robustness. Among the determinants, staking structure factors have the strongest influence, followed by brand-related and service quality factors, while economic factors exhibit the weakest impact. Overall, the study provides a data-driven perspective on staking behavior in PoS blockchain networks and offers practical implications for validator operation and delegation decisions.

KEY WORDS

Solana, Validator Staking, Machine Learning, Stacking Ensemble, SHAP

C3

Service Fee in C2C Secondhand Markets: Optimal Fee Design for Platform Governance Value

Guojun Ji, Chang Liu (School of Management, Xiamen University) and
Kim Hua Tan (Business School, Nottingham University)

Against the backdrop of the circular economy, the rapid growth of C2C secondhand markets has intensified concerns about transaction quality and platform governance. While most platforms rely on high commissions, Xianyu's shift from zero commission to extremely low tiered service fees is notable. We develop a model with a platform, consumers, and heterogeneous sellers (idle and professional) differing in fulfillment rates, where the platform maximizes governance value measured by successful transactions. Under a non-profit objective, the optimal base service fee is always zero, while tiered fees exhibit corner solutions (near 0% or 100%) depending on seller composition and reliability differences. Service fees improve governance only when professional sellers are few and unreliable; otherwise, a no-fee policy is preferable. Additionally, fees redistribute surplus from professional sellers to idle sellers and sometimes to consumers, offering insights into Xianyu's strategy and sustainable platform governance.

KEY WORDS

secondary market, service fee, platform governance, non-profit orientation

Service Fee in C2C Secondhand Markets: Optimal Fee Design for Platform Governance Value

Ji, G. J.[a], Liu, C.[a], and Tan, K. H.[b]*
[a] School of Management, Xiamen University
E-mail: jiking@xmu.edu.cn, liuchnag2022@stu.xmu.edu.cn
[b] Business School, Nottingham University
E-mail: kim.tan@nottingham.ac.uk

Abstract

Against the backdrop of the circular economy, the rapid growth of C2C secondhand markets has intensified concerns about transaction quality and platform governance. While most platforms rely on high commissions, Xianyu's shift from zero commission to extremely low tiered service fees is notable. We develop a model with a platform, consumers, and heterogeneous sellers (idle and professional) differing in fulfillment rates, where the platform maximizes governance value measured by successful transactions. Under a non-profit objective, the optimal base service fee is always zero, while tiered fees exhibit corner solutions (near 0% or 100%) depending on seller composition and reliability differences. Service fees improve governance only when professional sellers are few and unreliable; otherwise, a no-fee policy is preferable. Additionally, fees redistribute surplus from professional sellers to idle sellers and sometimes to consumers, offering insights into Xianyu's strategy and sustainable platform governance.

Keywords: secondary market, service fee, platform governance, non-profit orientation

Introduction

Globally, circular consumption models that promote sustainable development have been widely advocated and increasingly implemented (Van Loon et al. 2018). As a key component of sustainable supply chains and E-commerce, professional secondhand market platforms such as Xianyu, Zhuanzhuan, and Mercari have emerged with the development of the sharing economy and digital technology (Zhang et al. 2021). By substantially reducing information search and transaction costs, these platforms enhance the efficiency and attractiveness of idle goods circulation, thereby accelerating the growth of secondhand transactions (Liu et al. 2023). Market evidence supports this trend: approximately 62.9% of Americans sold at least one secondhand item online in the past year (Ritch et al. 2023), and the global secondary market is projected to double to a staggering $350 billion by 2027 (Chu 2024; Scarsella 2023). Secondhand transactions generate significant environmental benefits by extending product lifecycles and promoting the reuse of idle goods (Huang and Yue 2024). At the same time, digital platforms also facilitate the monetization of idle assets, lower the barriers to consumer participation, and stimulate economic activity through market expansion (Lo et al. 2019). Consequently, secondhand trading platforms have become a critical link in improving the circulation efficiency of consumer goods, enhancing social welfare, and advancing sustainable development (Tan et al. 2024).

At present, secondhand trading platforms mainly operate under three transaction structures: C2B2C (Consumer to Business to Consumer), C2C (Customer-to-Customer), and hybrid models. Under the C2B2C model, platforms such as Zhuanzhuan (China) and thredUP (United States) directly purchase used goods from consumers, conduct inspection and necessary refurbishment, and subsequently resell the products (Zeng et al. 2023). The C2C model, represented by Xianyu (China), Mercari (Japan), and

Poshmark (United States), relies on direct transactions between individual sellers and buyers; the platform only provides information services and matches needs, without being responsible for the transactions (Li et al. 2022). The hybrid model combines these two mechanisms, as exemplified by platforms like eBay. C2B2C secondhand platforms charge sellers commissions to cover operating costs, while C2C secondhand trading platforms have inconsistent fee models. Mercari and Poshmark both charge sellers commissions or transaction fees as a means of profit. However, despite its penetration rate of about 70% to 73% in China's secondhand e-commerce market, which is far higher than other competing platforms, Xianyu has not charged any fees to either party in the transaction from its establishment in 2014 to 2024. Therefore, when Xianyu announced that it would start charging sellers extremely low service fees from March 2024, this unusual move attracted widespread attention. This was not only because the platform suddenly switched from a free policy to a paid policy after 10 years of operation, but also because its service fee (maximum 1.6%) was far lower than other mainstream C2C platforms, Mercari: 10% (Mercari 2024) and Poshmark: 20% (Ben 2025).

Unlike most independently operated, profit-driven secondhand platforms, Xianyu does not aim for direct profit. Instead, its core objective is to promote the development of a circular economy by attracting new users and facilitating the circulation of idle goods. Therefore, Xianyu's C2C trading model has not charged sellers service fees for a considerable period. With the support of Taobao's massive traffic, this zero-fee strategy effectively lowers the barrier to entry for users participating in secondhand trading, significantly increasing platform activity (Thch 2024). However, this strategy also provides professional sellers with a near-rent-free traffic entry point. These professional sellers, by reselling large quantities of goods of unknown origin at low prices, have built arbitrage-oriented trading channels on Xianyu, leading to a surge in low-priced, low-quality goods on the platform. Idle sellers typically sell goods for their own use and provide relatively comprehensive information to assist buyers in their decisions. Their accounts are linked to personal payment and credit systems, resulting in lower overall fraud risk and relatively higher fulfillment rates. In contrast, professional sellers often engage in opportunistic behavior through false descriptions and malicious transactions (Hong 2015), leading to a sharp increase in disputes and complaints due to their low fulfillment rates . Clearly, the excessive expansion of professional sellers has not only lowered the platform's overall fulfillment rate but also damaged the platform's transaction order and trust foundation. Against this backdrop, Xianyu announced that it will curb the disorderly expansion of professional sellers by charging sellers tiered service fees, thereby upholding the platform's original purpose as a platform for personal secondhand transactions.

However, whether this strategic adjustment can effectively improve the platform ecosystem and promote the circulation of idle goods remains to be further tested. In the C2C model, introducing a service fee mechanism may inhibit professional sellers while discouraging idle seller participation and reducing transaction volume, thereby increasing prices and affecting consumer welfare. Furthermore, it is unclear whether the platform intends to gain direct economic benefits through this measure. If the platform's true motivation is to profit through service fees, then whether its fee design can achieve its expected goals remains to be explored. Clearly, Xianyu's service fee level is far lower than other mainstream C2C platforms. In a market environment where service fees are commonly used as a means to increase profits, Xianyu's strategic shift raises a real-world question that has not been fully explored: when a C2C secondhand trading platform is non-profit oriented, how should its optimal service fee strategy be designed? The impact of this service fee strategy on consumer and seller surplus in the secondhand market, as well as its differences from the service fee strategies of profit-oriented platforms, still lacks in-depth research. The literature relevant to this study focuses on three main areas: secondary markets, service fee strategy and platform governance. Current research on C2C secondhand trading platforms primarily focuses on profit maximization and pays limited attention to the role of secondhand consumption in supporting the circular economy. Secondly, previous studies have typically viewed service fees as a pricing tool for revenue generation, while there is a lack of research on using service fees as a governance tool to regulate participant behavior and enhance the platform's governance value.

To answer these questions, we constructed a model involving sellers, consumers, and a C2C secondhand platform. Unlike traditional studies that aim to maximize profits, we set the platform's objective as maximizing governance value. The platform sets tiered service fees to achieve this goal, with

governance value comprising two core parameters: average fulfillment rate and transaction volume. Under this model, we focus on analyzing the platform's optimal service fee strategy-whether or not to charge a service fee and how to set the tiered service fee structure. Furthermore, we discussed the impact of the platform's service fee strategy on sellers surplus and consumers surplus, and provided a comprehensive evaluation of the optimal service fee strategy.

Model

We consider a single period C2C secondhand platform that intermediates transactions between sellers and buyers. The platform may impose service fees as a governance tool to regulate market composition and transaction reliability. As shown in Figure 1.

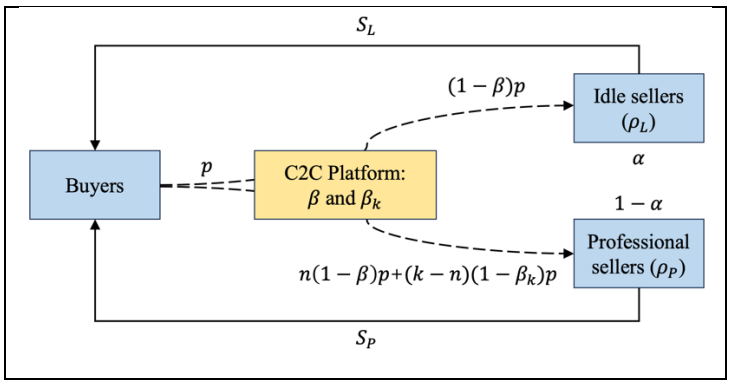


Figure 1. Diagram of C2C platform, sellers and buyers.

The assumptions of the model are as follows:

Sellers: Idle sellers supply at most one unit and typically sell self-used products. Professional sellers can supply multiple units (where k is an integer and $k > 1$) and engage in resale activities. The total mass of sellers is normalized to 1. Let $\alpha \in (0,1)$ and $1-\alpha$ represent the proportions of idle sellers and professional sellers, respectively. The supply from idle sellers and professional sellers is denoted as S_L and S_P, respectively.

Buyers: Buyers are heterogeneous in their baseline valuation for secondhand products. The value of a new product is denoted by v which is normalized to lie in $[0, 1]$. A secondhand product retains a fraction δ $(0 < \delta < 1)$ of its original value, so its intrinsic value is δv (Liu et al. 2018). Buyers observe the platform's service fee policy and understand that it affects seller participation and the overall trading environment.

Cost: Professional sellers face a per-unit reservation cost c, which captures the minimum compensation required for participating in resale activities, including sourcing effort, operational burden, and opportunity costs. The reservation cost is heterogeneous and follows a uniform distribution on $c \sim U[0,1]$(Huang and Yue 2024). Professional sellers act purely as resellers and derive no consumption value from the products they sell. Their participation depends on the resale margin.

Fulfillment rate: Empirical evidence suggests that sellers differ in fulfillment reliability (Yamamoto et al. 2019). Professional sellers typically operate with external supply chains and higher operational complexity, which increases the likelihood of misdescription, delayed delivery, or disputes. By contrast, idle sellers sell personally owned products with fewer transactions and lower uncertainty. Accordingly, we assume that idle sellers have a higher fulfillment rate than professional sellers: $\rho_L > \rho_P$, $\rho_L \in (0,1)$ and $\rho_P \in (0,1)$.

Service fees: In the benchmark model, the platform does not impose any service fees. Subsequently, in order to improve ecological governance, the platform introduced a tiered service fee system. Specifically, for each unit sold at price p, the platform charges a basic service fee at rate $\beta \in [0,1)$. For transactions beyond the n-th unit completed by a seller within a period, the service fee rate increases to $\beta_k \in (0,1)$, where $\beta_k > \beta$, n is an integer, $n \geq 2$ and $k > n$.

We therefore assume that buyers form rational expectations about the equilibrium average fulfillment rate induced by the service fee policy:

$$\overline{\rho} = \frac{S_L\rho_L + S_P\rho_P}{S_L + S_P}. \quad (1)$$

In the secondhand market, prices are determined endogenously by supply and demand (Xue et al. 2018; Yin et al. 2010). A buyer who pays price p obtains net utility $\mu = \overline{\rho}\delta v - p$. Each buyer purchases at most one unit if and only if $\mu > 0$. Platform governance aims to promote transactions involving idle goods and reduce fraud and disputes. Its value is reflected in higher transaction reliability and a greater presence of idle sellers. Drawing on literature from operations management regarding reliability-adjusted throughput (Li et al. 2023; Xue et al. 2018) and quality-adjusted output (Chen et al. 2016; Wang and Wright 2025), we define platform governance value as the number of successfully fulfilled transactions. This type of model structure appears in research on C2C platforms (Li et al. 2023) and the secondary market (Yin et al. 2010). The platform's governance value function is given by:

$$\Pi = \lambda\overline{\rho}Q. \quad (2)$$

Here, Q denotes the equilibrium transaction volume, and $\lambda > 0$ captures the platform's emphasis on governance. Table 1 summarizes the key parameters.

Table 1. Parameters and explanations

Parameters	Explanations	Parameters	Explanations
v	Value of the product	α	Proportion of idle sellers
p	Market clearing price	δ	Durability coefficient
c	Reservation cost	Π	Platform's governance value function
$\overline{\rho}$	Average fulfillment rate	λ	Platform's emphasis in ecological
μ	The utility of buyers	k	Starting quantity for tiered service rate
β	Basic service fee rate	ρ_L	Fulfillment rate of idle sellers
β_k	Tiered service fee rate	ρ_P	Fulfillment rate of professional sellers

Equilibrium result

Benchmark model

In the benchmark model, $\beta = \beta_k = 0$. The total demand from buyers is D_0. On the supply side, the utility of idle sellers and professional sellers are $U_L = p - \delta v$ and $U_P = k(p - c)$, respectively. The participation conditions for the two types of sellers are as follows: $U_L > 0$ and $U_P > 0$.By supply and demand equilibrium, $S_0 = S_L + S_P = D_0$, we get $p_0^* = \frac{1}{\frac{\alpha}{\delta} + k(1-\alpha) + \frac{1}{\overline{\rho}\delta}}$, $Q_0^* = 1 - \frac{p_0^*}{\overline{\rho}\delta}$ and $\Pi_0^* = \lambda\overline{\rho}_0^* Q_0^*$.

Tiered service fees (F)

The platform introduces tiered service fees as a governance tool. The platform first determines the service fee rates, after which sellers decide whether to participate based on their individual costs and benefits. For each unit sold, idle sellers and the first n units sold by professional sellers are charged basic service fee β. For additional units sold by professional sellers, the tiered service fee increases to β_k. Idle sellers' and Professional sellers' utility becomes $U_L^F = (1-\beta)p - \delta v$ and $U_P^F = n(1-\beta)p + (k-n)(1-\beta_k)p - kc$, respectively. By supply and demand equilibrium $S^F = S_L^F + S_P^F = D_0$, we get the equilibrium price $p_F^* = \frac{1}{\frac{1}{\overline{\rho}\delta} + \frac{\alpha(1-\beta)}{\delta} + (1-\alpha)[k(1-\beta_k) + n(\beta_k - \beta)]}$, the equilibrium transaction volume $Q_F^* = 1 - \frac{p_F^*}{\overline{\rho}\delta}$, and the platform's governance value function $\Pi_F^* = \lambda\overline{\rho}_F^* Q_F^*$. The optimal β and β_k are shown in Table 2, which represents the optimal strategies for the platform under different conditions.

Table 2. The optimal β and β_k.

	$\alpha(\rho_L-\rho_P)<\frac{T}{2+T}$	$\frac{T}{2+T}<\alpha(\rho_L-\rho_P)<\frac{T+A}{2+T+A}$	$\alpha(\rho_L-\rho_P)>\frac{T+A}{2+T+A}$
$\frac{\rho_L}{\rho_P}<2+\frac{(1-\alpha)n\delta}{\alpha}$	$\beta^*=0, \beta_k^*=0^+$	$\beta^*=0, \beta_k^*=\underset{\beta_k\to 0^+ \text{ or } 1^-}{\arg\max}\ \Pi(\beta_k)^*$	$\beta^*=0, \beta_k^*=1^-$
$2+\frac{(1-\alpha)n\delta}{\alpha}<\frac{\rho_L}{\rho_P}<2+\frac{(1-\alpha)k\delta}{\alpha}$	None	$\beta^*=0, \beta_k^*=\underset{\beta_k\to 0^+ \text{ or } 1^-}{\arg\max}\ \Pi(\beta_k)^*$	$\beta^*=0, \beta_k^*=1^-$
$\frac{\rho_L}{\rho_P}>2+\frac{(1-\alpha)k\delta}{\alpha}$	None	None	$\beta^*=0, \beta_k^*=1^-$

$T=\alpha\rho_L+n(1-\alpha)\delta\rho_P$ and $A=(k-n)(1-\alpha)\delta\rho_P$.

Analysis and comparison

This section will compare and analyze the results, supplemented by numerical studies for verification.

Seller's strategy

Proposition 1 The impact of service fees on sellers' decisions:

$$(1)\ \frac{\partial s_L^F}{\partial\beta}<0, \frac{\partial s_P^F}{\partial\beta}<0, \frac{\partial(\frac{s_L^F}{s_L^F+s_P^F})}{\partial\beta}<0;\ (2)\ \frac{\partial s_L^F}{\partial\beta_k}=0, \frac{\partial s_P^F}{\partial\beta_k}<0, \frac{\partial(\frac{s_L^F}{s_L^F+s_P^F})}{\partial\beta_k}>0.$$

An increase in the basic service fee reduces the revenues of both idle sellers and professional sellers. As a result, higher basic service fee induces gradual exit by idle sellers and a reduction in professional sellers' participation, leading to a decline in total supply. Because each idle seller contributes only one unit, their exit has a disproportionately large effect on overall supply, reducing the share of idle seller supply on the platform. In contrast, an increase in the tiered service fee does not affect idle sellers but constrains professional sellers, who adjust their supply to avoid incurring higher fees. Consequently, higher tiered service fees increase the relative proportion of idle seller supply on the platform.

Platform's strategy

Proposition 2. The optimal strategy for maximizing platform governance value is as follows:

If $\alpha(\rho_L-\rho_P)<\frac{T(T+A)}{T+(T+1)(T+A)}$, $\beta^*=0$ and $\beta_k^*=0^+$, $p_F^*>p_0^*, \overline{\rho}_F^*>\overline{\rho}_0, Q_F^*<Q_0^*, \Pi_F^*<\Pi_0^*$;

If $\alpha(\rho_L-\rho_P)>\frac{T(T+A)}{T+(T+1)(T+A)}$, $\beta^*=0$ and $\beta_k^*=1^-$, $p_F^*>p_0^*, \overline{\rho}_F^*>\overline{\rho}_0, Q_F^*<Q_0^*, \Pi_F^*>\Pi_0^*$.

Proposition 2 shows that the optimal basic service fee is $\beta^*=0$, while maximizing governance value requires selecting β_k^* based on the trade-off between the average fulfillment rate and transaction volume.

When $\alpha(\rho_L-\rho_P)$ is low, the share of idle sellers is moderate and the difference in fulfillment rates between the two types of sellers is small. Since professional sellers are essential for sustaining transaction volume, tiered service fees reduce their participation and thus decrease total supply. Although this improves the average fulfillment rate and increases prices, the resulting governance gains are insufficient to offset the loss from reduced transaction volume, leading to a governance value that is slightly lower than that under no service fees. In this case, the platform should maintain an extremely low tiered service fee, a counterintuitive yet reasonable result that provides theoretical support for the low-fee strategy observed on Xianyu.

In contrast, when $\alpha(\rho_L-\rho_P)$ is high, idle sellers dominate the market and the difference in fulfillment rates is substantial. Professional sellers are few but exhibit low reliability, which harms both consumers and idle sellers. In this setting, improving fulfillment reliability becomes more important than maintaining transaction volume. The platform should therefore impose a high tiered service fee to restrain professional sellers and induce their exit. The resulting improvement in fulfillment reliability generates governance gains that exceed the loss from reduced transaction volume, so that governance value becomes higher than in the no-fee scenario.

Figure 2 presents numerical results showing how the equilibrium result varies with β_k.

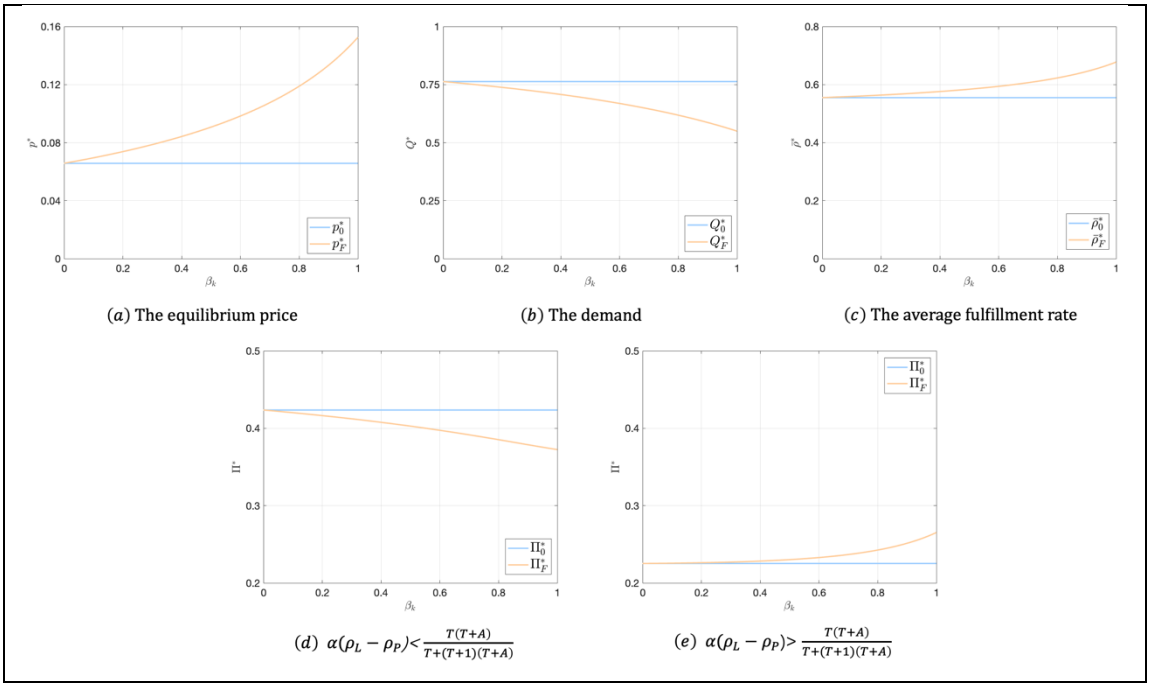


Figure 2. Equilibrium results with β_k. ($\alpha = 0.8, \delta = 0.5, \rho_L = 0.9, \rho_P = 0.5, \lambda = 1.$)

Surplus analysis

This section analyzes consumer surplus (CS), idle sellers surplus (LS), and professional sellers surplus (PS) under different service fee strategies. Figure 3 presents numerical results showing how the surplus varies with β_k.

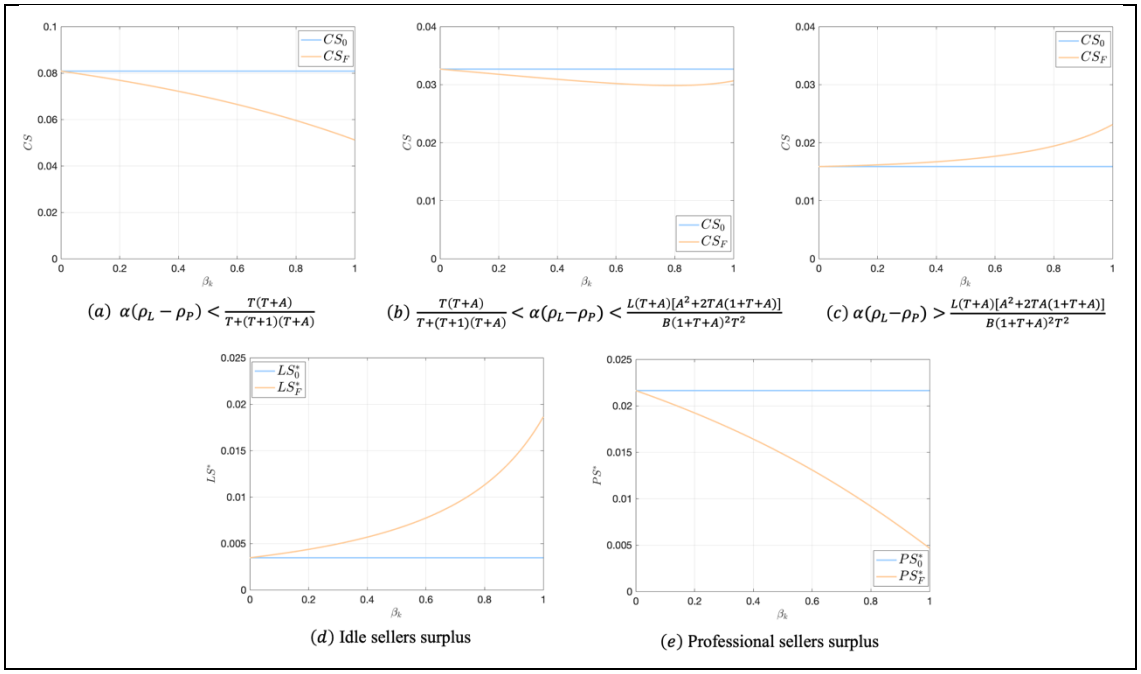


Figure 3. Surplus with β_k. ($\alpha = 0.8, \delta = 0.5, \rho_L = 0.9, \rho_P = 0.5$ and $\lambda = 1.$)

Proposition 3.

(1) When $\alpha(\rho_L - \rho_P) < \frac{T(T+A)}{T+(T+1)(T+A)}$, $CS_F < CS_0$, $LS_F > LS_0$, $PS_F < PS_0$.

(2) When $\frac{T(T+A)}{T+(T+1)(T+A)} < \alpha(\rho_L - \rho_P) < \frac{L(T+A)[A^2+2TA(1+T+A)]}{B(1+T+A)^2T^2}$, $CS_F < CS_0$, $LS_F > LS_0$, $PS_F < PS_0$.

(3) When $\alpha(\rho_L - \rho_P) > \frac{L(T+A)[A^2+2TA(1+T+A)]}{B(1+T+A)^2T^2}$, $CS_F > CS_0$, $LS_F > LS_0$, $PS_F < PS_0$.

According to Proposition 3, the platform's service fee strategy effectively reallocates surplus from low-quality professional sellers to idle sellers and consumers, with priority given to idle sellers. This highlights the platform's objective of protecting idle sellers and supporting the circular economy. When $\alpha(\rho_L - \rho_P)$ is low, refraining from setting a service fee is the optimal strategy for the platform. Conversely, when $\alpha(\rho_L - \rho_P)$ is high, implementing a higher tiered service fee is necessary to enhance governance and protect consumer welfare.

Conclusions

This study is motivated by the rapid expansion of the secondhand market under the circular economy and the governance challenges faced by C2C platforms: how should a non-profit oriented platform design its service fee strategy to balance transaction efficiency and governance outcomes? The main findings reveal several important and non-obvious insights. First, the optimal basic service fee is zero, and the optimal tiered service fee is corner-based, depending on seller composition and the fulfillment rate gap. Notably, when the market is relatively healthy, the optimal strategy is to maintain an extremely low or even negligible service fee. This implies that Xianyu's long-standing low service fee policy is not merely a competitive choice, but can be interpreted as an optimal governance outcome. Second, service fees do not necessarily improve platform performance; when the fulfillment rate difference is limited, introducing service fees may reduce governance value due to the contraction of transaction volume. Third, only when the fulfillment rate gap becomes sufficiently large does a high tiered service fee become optimal, as it effectively disciplines low-quality professional sellers and improves overall market reliability.

Based on the conclusions of this study, we draw the following management implications:

(1) Platform governance should be contingent on market conditions rather than uniform pricing policies. When the market is relatively healthy, introducing service fees may be counterproductive. The long-term zero fee policy of Xianyu illustrates this point: when idle sellers dominated and fulfillment reliability was high, maintaining zero fees sustained participation and transaction volume. In this context, non-intervention can represent an optimal governance strategy. (2) Service fees should be designed as governance tools to shape seller composition rather than as revenue instruments. A zero or low basic service fee sustains idle seller participation, while tiered service fees can selectively regulate professional sellers. The recent policy of Xianyu includes a basic fee of about 0.6% and a maximum tiered rate of 1.6%, which targets high-volume sellers while minimizing the impact on ordinary users. (3) Effective platform governance requires the integration of pricing mechanisms with complementary governance tools. Measures such as authentication services, reputation systems, and transaction security mechanisms play a critical role in reducing information asymmetry and improving transaction reliability. The costs of these governance tools can be covered through service fees, ensuring their sustained implementation. At the same time, service fees signal the presence of governance efforts to consumers, thereby enhancing trust and facilitating the circulation of idle resources. (4) Governance strategies should evolve with market development and balance multiple objectives. As platforms mature and professional sellers expand, governance must shift from market expansion toward quality control and ecosystem stability. Platforms should balance transaction volume, fulfillment reliability, and user experience to enhance long-term governance value, strengthen reputation, and improve user retention.

References

Ben. 2025. "Poshmark Vs. Mercari Vs. Ebay: Which Is Better for Reselling Clothes?", from https://www.easyship.com/blog/mercari-vs-ebay-vs-poshmark

Chen, J., Fan, M., and Li, M. 2016. "Advertising Versus Brokerage Model for Online Trading Platforms," MIS Quarterly (40:3), pp. 575-596.

Chu, H. 2024. "Mitigating Consumer Guilt through Resale," Marketing Intelligence & Planning (42:3), pp. 515-533.

Hong, I. B. 2015. "Understanding the Consumer's Online Merchant Selection Process: The Roles of Product Involvement, Perceived Risk, and Trust Expectation," International journal of information management (35:3), pp. 322-336.

Huang, M., and Yue, H. 2024. "Service Strategies for Second-Hand C2c Platforms: Inspection Service Option," Electronic Commerce Research, pp. 1-29.

Li, L., Fang, X., and Lim, Y. F. 2023. "Asymmetric Information of Product Authenticity on C2C E-Commerce Platforms: How Can Inspection Services Help?," Manufacturing & Service Operations Management (25:2), pp. 631-647.

Li, Z., Dai, R., Feng, X., and Xiong, Y. 2022. "The Analysis of Two-Way E-Commerce Credit Evaluation Model Based on the C2c Mode," Journal of Global Information Management (JGIM) (30:11), pp. 1-21.

Liu, J., Zhai, X., and Chen, L. 2018. "The Interaction between Product Rollover Strategy and Pricing Scheme," International Journal of Production Economics (201), pp. 116-135.

Liu, Y., Wan, Y., and Kang, J. 2023. "Impact of Community-Based Governance Mechanisms on Transaction Intention on a Second-Hand Trading Platform," Journal of Theoretical and Applied Electronic Commerce Research (18:1), pp. 689-705.

Lo, C. J., Tsarenko, Y., and Tojib, D. 2019. "To Tell or Not to Tell? The Roles of Perceived Norms and Self-Consciousness in Understanding Consumers' Willingness to Recommend Online Secondhand Apparel Shopping," Psychology & Marketing (36:4), pp. 287-304.

Mercari. 2024. "Fees on Mercari (Help Center)." from https://www.mercari.com/us/help_center/article/169/

Ritch, E. L., Canning, C., and Siddiqui, N. Q. 2023. "Resell–Rewear: Examining Consumer Behaviour to Advance the Fashion Circular Economy," in Pioneering New Perspectives in the Fashion Industry: Disruption, Diversity and Sustainable Innovation. Emerald Publishing Limited, pp. 173-184.

Scarsella, A. 2023. " Worldwide Used Smartphone Forecast, 2022–2026." from https://www.businesswire.com/

Tan, L., Yang, Z., Irfan, M., Ding, C. J., Hu, M., and Hu, J. 2024. "Toward Low-Carbon Sustainable Development: Exploring the Impact of Digital Economy Development and Industrial Restructuring," Business Strategy and the Environment (33:3), pp. 2159-2172.

Thch, P. 2024. "Saying Goodbye to the Free Era, Is Xianyu (a Second-Hand Marketplace Platform) Eager to Make Money?", from https://www.jiemian.com/article/11525381.html

Van Loon, P., Delagarde, C., and Van Wassenhove, L. N. 2018. "The Role of Second-Hand Markets in Circular Business: A Simple Model for Leasing Versus Selling Consumer Products," International Journal of Production Research (56:1-2), pp. 960-973.

Wang, C., and Wright, J. 2025. "Regulating Platform Fees," Journal of the European Economic Association (23:2), pp. 746-783.

Xue, Y., Caliskan-Demirag, O., Chen, Y. F., and Yu, Y. 2018. "Supporting Customers to Sell Used Goods: Profitability and Environmental Implications," International Journal of Production Economics (206), pp. 220-232.

Yamamoto, H., Sugiyama, N., Toriumi, F., Kashida, H., and Yamaguchi, T. 2019. "Angels or Demons? Classifying Desirable Heavy Users and Undesirable Power Sellers in Online C2c Marketplace," Journal of Computational Social Science (2:2), pp. 315-329.

Yin, S., Ray, S., Gurnani, H., and Animesh, A. 2010. "Durable Products with Multiple Used Goods Markets: Product Upgrade and Retail Pricing Implications," Marketing Science (29:3), pp. 540-560.

Zeng, Y., Bai, X., and Wang, Y. 2023. "How Should Resale Platforms Operate the Customer-to-Business-to-Customer Mode?," Electronic Commerce Research and Applications (58), p. 101251.

Zhang, Y., Huang, M., Tian, L., Jin, D., and Cai, G. G. 2021. "Build or Join a Sharing Platform? The Choice of Manufacturer's Sharing Mode," International Journal of Production Economics (231), p. 107811.

C3

How AI Credit Scoring Influences Consumer Satisfaction and Purchase Intention in AI-Driven E-Commerce: Evidence from Blibli PayLater

Sheila Silvia Permatasari, Syuan Yi Jiang (National Cheng Kung University).

The integration of Artificial Intelligence (AI) and Financial Technology (FinTech) within e-commerce platforms represents a critical evolution in the AI Transformation (AX) ecosystem. This research-in-progress investigates the behavioral responses of consumers to AI-driven credit scoring systems, specifically focusing on the partnership between Blibli and IndoDana Fintech in Indonesia. By analyzing consumer perceptions of the AI's speed, fairness, and transparency, this study proposes a conceptual model to evaluate how AI credit scoring influences consumer satisfaction and subsequent purchase intention. Utilizing a quantitative approach via Partial Least Squares Structural Equation Modeling (PLS-SEM), the anticipated findings aim to demonstrate that AI credit scoring is not merely a backend risk-mitigation tool, but a primary driver of the frontend customer experience.

KEY WORDS

Artificial Intelligence, FinTech Lending, E-commerce, credit scoring

How AI Credit Scoring Influences Consumer Satisfaction and Purchase Intention in AI-Driven E-Commerce: Evidence from Blibli PayLater

Permatasari, S. S.[a] *and Jiang, S. Y.*[b]
[a] *Department of Business Administration, National Cheng Kung University*
E-mail: r68137010@gs.ncku.edu.tw
[b] *Department of Business Administration, National Cheng Kung University*
E-mail: syuan.yi.jiang@gs.ncku.edu.tw

Abstract

The integration of Artificial Intelligence (AI) and Financial Technology (FinTech) within e-commerce platforms represents a critical evolution in the AI Transformation (AX) ecosystem. This research-in-progress investigates the behavioral responses of consumers to AI-driven credit scoring systems, specifically focusing on the partnership between Blibli and IndoDana Fintech in Indonesia. Underpinned by the Stimulus-Organism-Response (S-O-R) framework, this study proposes a conceptual model where AI-driven attributes act as environmental stimuli that shape the consumer's internal psychological state (satisfaction), ultimately leading to a behavioral response (purchase intention). Utilizing a quantitative approach via Partial Least Squares Structural Equation Modeling (PLS-SEM), the anticipated findings aim to demonstrate that AI credit scoring is not merely a backend risk-mitigation tool, but a primary driver of the frontend customer experience.

Keywords: Artificial Intelligence, FinTech lending, e-commerce, credit scoring, algorithmic experience (AX)

Introduction

The digital economy in Indonesia is experiencing a rapid convergence of e-commerce and FinTech lending. A prominent example of this synergy is the "Buy Now, Pay Later" (BNPL) feature, exemplified by the partnership between the e-commerce giant Blibli and the FinTech lender IndoDana. To manage risk seamlessly during the checkout process, these platforms employ Artificial Intelligence (AI) to conduct instant credit scoring (Aisjah, 2024).

Credit scoring is a statistical analysis utilized by financial institutions to assess the creditworthiness of prospective borrowers based on historical data. The AI system aggregates data from traditional sources, such as the Financial Information Services System (SLIK) managed by Indonesia's Financial Services Authority (OJK), alongside application data and digital behavioral footprints to bypass traditional logistic regression models in favor of more dynamic neural networks.

While the operational benefits of AI credit scoring, such as accelerated approval decisions, reduced human subjectivity, and democratized credit access are well documented (Shukla et al., 2025), there is a distinct research gap regarding the user's psychological and behavioral response to being evaluated by an algorithm. Consumers often experience "lack of transparency anxiety" when assessed by non-human agents (Lopez and Garza, 2023). This study aims to address this gap by utilizing the Stimulus-Organism-Response (S-O-R) framework to map how the perceived quality of the AI (speed, fairness, transparency, and ease of use) impacts user satisfaction and their willingness to utilize the PayLater

service for future purchases (Vaccaro and Karahalios, 2020). Based on the introduction, the following research questions are directly derived from the aims and hypotheses presented in this paper:

1. How does the perceived quality of AI credit scoring (speed, fairness, transparency, and ease of use) affect consumer satisfaction in e-commerce?
2. How does consumer satisfaction affect a user's intention to purchase using AI-driven PayLater services?
3. What is the direct impact of the perceived quality of AI credit scoring on consumer purchase intention?
4. To what extent does consumer satisfaction mediate the relationship between AI credit scoring and purchase intention?

Background

This study adopts the Stimulus-Organism-Response (S-O-R) framework to explain the mechanisms behind consumer behavior within the AI credit scoring ecosystem. The S-O-R theory posits that environmental cues (Stimulus) influence an individual's internal state (Organism), which in turn drives a specific behavioral outcome (Response). By clarifying the multi-dimensional constructs of AI credit scoring through this lens, this study expands the theoretical contribution of Algorithmic Experience (AX) literature within emerging market FinTech.

Stimulus (S) - AI Credit Scoring: In this context, the stimulus consists of the attributes of the AI Credit Scoring system. The technical and ethical quality of the system, operationalized through processing speed, algorithmic fairness, process transparency, and ease of use, serves as the digital environmental cue received by the user. Speed is the velocity of real-time data processing and decision-making. Fairness is the perceived absence of algorithmic bias in credit limit allocation. Transparency is the clarity of the "black box" logic and data usage. Ease of Use is the seamlessness of the digital interface during assessment.

Organism (O) - Consumer Satisfaction: The internal state is represented by Consumer Satisfaction. This is the psychological and affective evaluation of the user after interacting with the AI stimulus, which helps mitigate algorithmic anxiety.

Response (R) - Purchase Intention: The behavioral outcome is represented by Purchase Intention. This is the user's commitment or willingness to execute a transaction on the Blibli platform using the PayLater facility.

AI Credit Scoring on Consumer Satisfaction

When users perceive the algorithmic assessment (Stimulus) as fast, transparent, fair, and easy to navigate, their cognitive effort is reduced, and their trust in the platform increases, leading to higher satisfaction (Organism) (Afroogh et al., 2025).

H1: Perceived quality of AI Credit Scoring (Stimulus) has a positive effect on Consumer Satisfaction (Organism).

Consumer Satisfaction on Purchase Intention

Satisfied users who experience a frictionless credit approval process are highly likely to reuse BNPL services, increase their basket size, and develop long-term loyalty toward the Blibli platform (Hegawan et al., 2024).

H2: Consumer Satisfaction (Organism) has a positive effect on Purchase Intention (Response).

AI Credit Scoring on Purchase Intention

The perceived efficiency of the AI removes immediate financial barriers, acting as a direct catalyst for purchase intention by providing high financial self-efficacy at the point of sale (Simiyu et al. 2025).

H3: Perceived quality of AI Credit Scoring (Stimulus) has a direct positive effect on Purchase Intention (Response).

The Mediating Role of Consumer Satisfaction

While AI scoring provides the means to purchase, satisfaction provides the motivation. This study posits that consumer satisfaction acts as the crucial psychological bridge within the S-O-R flow.

H4: Consumer Satisfaction (Organism) mediates the relationship between AI Credit Scoring (Stimulus) and Purchase Intention (Response).

Methodology

This research-in-progress employs a quantitative, cross-sectional design to test the proposed S-O-R hypotheses. AI Risk Assessment Model in this study utilizes a hybrid of traditional data (SLIK/OJK) and non-traditional digital behavioral footprints. The AI employs deep learning neural networks to assess creditworthiness instantly at the point of sale.

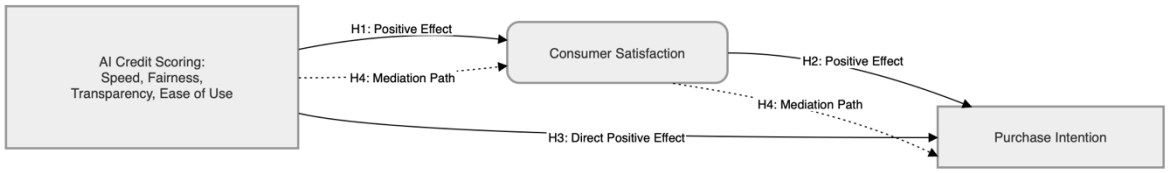


Figure 1. Research Model

Participants and Data Collection

The target population comprises Indonesian e-commerce consumers who have applied for or utilized the Blibli PayLater feature powered by IndoDana. A sample size of 150–200 respondents will be recruited using purposive sampling. To strengthen empirical validation prior to full deployment, a pilot study involving 30 participants will be conducted to ensure the survey instruments, adapted from validated scales in prior literature, demonstrate adequate reliability and discriminant validity using a 5-point Likert scale.

Analytical Approach and Positionality

Data cleaning and descriptive statistics will be conducted using SPSS. To test the complex relationships and mediating effects within the conceptual model, Partial Least Squares Structural Equation Modeling (PLS-SEM) will be executed using SmartPLS (Hair et al., 2022). The PLS-SEM approach will rigorously validate the model in two stages: first, by assessing the measurement model (outer model) for construct reliability and validity; and second, by evaluating the structural model (inner model) using bootstrapping to determine the significance of path coefficients and the strength of the mediation effect. The measurement model includes Cronbach's Alpa, AVE, and HTMT for discriminant validity. Meanwhile, the structural model includes bootstrapping for path siginificance.

Expected Results

As early-stage research, the anticipated results are expected to validate the proposed S-O-R structural model in the context of Algorithmic Experience (AX). First, the measurement model is expected to confirm the construct clarity of the multi-dimensional AI stimulus (fairness, speed, transparency, and ease of use). Regarding hypothesis testing, the most critical expected finding is the confirmation of H4, establishing Consumer Satisfaction (O) as a strong mediator. Finally, it is expected to demonstrate that while the AI stimulus directly enables purchases (H3), sustainable motivation for long-term e-commerce metrics is realized only when the technical stimulus is capable of generating a positive psychological state (satisfaction) within the user (H1 and H2).

This study expands AX literature by proving that algorithmic quality is not just a backend function but a crucial environmental stimulus for customer relationship management in emerging markets.

Conclusion

By employing the S-O-R theory, this study advances the theoretical contribution to AI Transformation (AX) in e-commerce by explaining why algorithmic quality impacts economic behavior, moving beyond mere conceptualization to empirical validation. The early-stage findings are expected to show that transparent and fast AI credit scoring is not merely a backend risk tool, but a crucial environmental stimulus integral to customer relationship management in the digital era. Limitations of this proposed research include its geographical focus on Indonesia and reliance on a single platform ecosystem (Blibli). Future work will involve executing the full data collection phase and comparing behavioral responses across platforms utilizing different AI scoring models.

References

Afroogh, S., Akbari, A., Malone, E. et al. 2024. "Trust in AI: progress, challenges, and future directions." *Humanities and Social Sciences Communications* (11:1568), pp 1-30. https://doi.org/10.1057/s41599-024-04044-8

Aisjah, S. 2024. "Intention to Use Buy-Now-Pay-Later Payment System Among University Students: A Combination of Financial Parenting, Financial Self-Efficacy, and Social Media Intensity," *Cogent Social Sciences* (10:1), pp. 1-18. https://doi.org/10.1080/23311886.2024.2306705

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. 2022. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.), Thousand Oaks, CA: Sage Publications. https://www.researchgate.net/publication/354331182_A_Primer_on_Partial_Least_Squares_Structural_Equation_Modeling_PLS-SEM

Hegawan, C. I. P., Faustine, T. M., Wijayanti, S., and Manaf, P. A. 2024. "Exploring the Influence of Perceived Usefulness and Value of PayLater/BNPL on Satisfaction, Impulsive Buying, and Post-Purchase Intention," *International Journal of Emerging Trends in Business, Economics* (3:2), pp. 45-60. https://doi.org/10.55217/103.v15i1.667

Lopez, A., and Garza, R. 2023. "Consumer Bias Against Evaluations Received by Artificial Intelligence: The Mediation Effect of Lack of Transparency Anxiety," *Journal of Research in Interactive Marketing* (17:6), November, pp. 831-847. https://doi.org/10.1108/JRIM-07-2021-0192

Shukla, S. K., Singh, R., Agrawal, S., Dwivedi, A. 2025. "Artificial Intelligence and the Future of Credit Scoring in FinTech: A Paradigm Shift," *International Journal of All Research Education & Scientific Methods* (13:2), February, pp. 904-909. https://doi.org/10.56025/IJARESM.2025.130225904

Simiyu, N. N., Kumar, S., and Kivuyo, R. B. 2025. "Exploring the Influence of Financial Self-Efficacy and Facilitating Conditions on Intentions and Actual Borrowing in Buy Now, Pay Later Services," *International Journal of Bank Marketing* (43:8), pp. 1705-1731. https://doi.org/10.1108/IJBM-09-2024-0527

Vaccaro, K., and Karahalios, K. 2020. "'At the End of the Day Facebook Does What It Wants': How Users Experience Contesting Algorithmic Content Moderation," *Proceedings of the ACM on Human-Computer Interaction* (4:CSCW2), October, pp. 1-22. https://s3.amazonaws.com/kvaccaro.com/documents/vaccaro_cscw2020.pdf

C3

Does the Adoption of Digital Wealth Management Platforms (DWPs) Result in Positive Net Benefits? Retail Investors' Perspective on DWP Success

Yong Yin, Kyung Young Lee, Bo Yu, Colin Conrad, Thomas Kb Koo (Dalhousie University).

The rising popularity of digital wealth management platforms (DWPs) has transformed how retail investors manage their portfolios. While DWPs provide easy access to advanced financial tools, their benefits for individual investors remain underexplored. This study addresses this gap by examining key factors influencing the successful adoption of DWPs and their impacts, using an integrated framework that combines the information systems success model, the service–profit chain model, and the behavioral investor types framework. Data were collected from 250 global DWP users. The research model evaluated how three quality dimensions—service, system, and information quality—affect system use and user satisfaction. Post-hoc multigroup analysis explored how behavioral investor types affect outcomes across different investor profiles. Findings show that service and information quality significantly impact system use and user satisfaction, while system quality positively affects user satisfaction but not system use. Both system use and user satisfaction strongly contribute to net benefits. Additionally, customer loyalty enhances the impact of service quality on engagement with the platform. This study broadens the information systems success model by integrating behavioral finance insights, highlighting how investor behavior influences DWP adoption and effectiveness. Practical implications suggest that DWP providers should personalize services to meet diverse investor needs.

KEY WORDS

Digital wealth management, IS success model, Behavioral investor types, Net benefits, IS quality

Does the Adoption of Digital Wealth Management Platforms (DWPs) Result in Positive Net Benefits? Retail Investors' Perspective on DWP Success

Yin, Y., Lee, K.Y., and Yu B.[a] *Conrad, C.*[b]*, and Koo T.K.B.*[c]
[a] *Department of Management Science and Information Systems, Dalhousie University*
E-mails: yn431051@dal.ca, bo.yu@dal.ca, kyunglee@dal.ca
[b] *Department of Information Science, Dalhousie University*
E-mail: colin.conrad@dal.ca
[c] *Department of Marketing, Dalhousie University*
E-mail: tom.kb.koo@dal.ca

Extended Abstract

Digital wealth management platforms (DWPs) have transformed financial services by offering streamlined, algorithm-driven solutions for wealth management. According to Deloitte (2020), DWPs are defined as digitally enabled channels that offer services such as portfolio management, account aggregation, reporting, trading capabilities, and access to research and investment ideas. The rise of Robo-advisors, such as Betterment in the U.S., exemplifies this shift, making investment services more accessible, customized, and convenient for both investors and wealth managers. While many DWPs are offered by digital brokers such as Wealthsimple, Coinbase, and Easy Securities, which provide diversified investment products, DWPs in the present research also include conventional mobile banking apps, browser-based banking portals, and telephone banking, as long as they offer the aforementioned services. DWPs deliver substantial benefits, including cost reduction, 24/7 accessibility, and personalized strategies (Boreiko and Massarotti, 2020). They democratize financial services, allowing investors of all backgrounds to access wealth management and optimize their portfolios through data-driven recommendations. Despite these advantages, the factors influencing DWP adoption and their benefits for individual investors remain underexplored.

Previous studies have examined the drivers and impacts of DWP adoption and its performance outcomes, but often without considering the combination of user behavior traits and their net benefits, which are essential for understanding the system's ultimate success, according to the information systems (IS) success model (DeLone and McLean, 2003). The model highlights six key success dimensions: system quality, information quality, service quality, system use, user satisfaction, and net benefits. However, few studies have applied this model to the broader wealth management sector, particularly in the context of DWPs. By incorporating the service-profit chain model (Heskett *et al*, 1994), which emphasizes the importance of customer loyalty in driving business success, and the behavioral investor types (BITs) framework (Pompian, 2008), which categorizes investors based on behavioral traits (Passive Preserver, Friendly Follower, Independent individualist, and Active Accumulator), this study seeks to provide a comprehensive analysis of the drivers of DWP adoption and its benefits for individual users. Specifically, the primary objectives of this research are threefold. First, it examines how service quality, system quality, and information quality affect system use and user satisfaction. Second, it explores whether system use and user satisfaction contribute to net benefits for individual investors. Third, it investigates whether customer loyalty positively moderates the relationship between service quality and system use, as well as between service quality and user satisfaction. Lastly, the moderating effect of BITs on the relationship between system use and net benefits and between user satisfaction and net benefits is analyzed. In sum, our research model evaluated how three quality dimensions, service, system, and information quality, affect system use and user satisfaction. Post-hoc multigroup analysis explored how behavioral investor types affect outcomes across different investor profiles. The following are the research hypotheses.

H1a-b: The service quality of DWPs is positively related to system use (a) and user satisfaction (b).

H2a-b: The system quality of DWPs is positively related to system use (a) and user satisfaction (b).

H3a-b: The information quality of DWPs is positively related to system use (a) and user satisfaction (b).

H4: System use of DWPs is positively related to net benefits.

H5: User satisfaction with DWPs is positively related to net benefits.

H6: The system use of DWPs and user satisfaction are positively associated.

H7a-b: Customer loyalty positively moderates the relationship between the service quality of DWPs and their system use (a) and user satisfaction (b).

H8a-b: BITs moderate the relationship between DWPs' system use (a) / user satisfaction (b) and net benefits, such that more aggressive investors will exhibit a stronger relationship between system use and net benefit.

Data was collected from 250 global DWP users, using a survey administered on Prolific. Eligibility criteria included being at least 18 years old, having a minimum of three months' experience using DWPs for investment purposes (e.g., fixed income, mutual funds, stocks, cryptocurrencies), and an intent to invest at least USD 500. The three-month minimum ensured familiarity with the platform's interface and features, while the USD 500 threshold aligned with standard account requirements, helping ensure relevant and informed responses. For data analysis, partial least squares structural equation modeling (PLS-SEM) was used.

Eight out of nine hypotheses for the direct relationships (H1-H6) were supported, with H2a as the exception; there was no statistically significant association between system quality and system use ($\beta = -0.073$, $t = 1.089$). A strong correlation between system use and user satisfaction was confirmed by a two-tailed Pearson test at the 0.001 significance level ($\beta = 0.27$, $t = 4.417$). Among the nine control variables, only occupation showed a significant relationship with net benefits at the 0.05 level. Individuals in high-income professions, such as business owners and senior managers, perceive greater benefits from DWPs than those in lower-income groups, likely due to diversified income sources like dividends and interest. The model explained 46.2% of the variance in net benefits. Figure 1 illustrates the research model and the results.

For H7 and H8, only the moderating effect of customer loyalty on the relationship between service quality and system use was significant, supporting H7a, while H7b, H8a, and H8b were not supported (Table 1). However, a post-hoc multigroup analysis divided respondents into 'less-aggressive' (passive preservers and friendly followers: 98 respondents) and 'more-aggressive' (independent individualists and active accumulators: 152 respondents) groups. This multigroup analysis reveals distinct behaviors within these broader BIT classifications. The results (Table 2) highlight a strong relationship between system use and net benefits for the more aggressive group, indicating a stronger association between DWP use and net benefits for independent individualists and active accumulators. Additionally, the relationship between user satisfaction and net benefits was statistically significant for both groups at the 0.001 level, underscoring that user satisfaction promotes net benefits across all BITs.

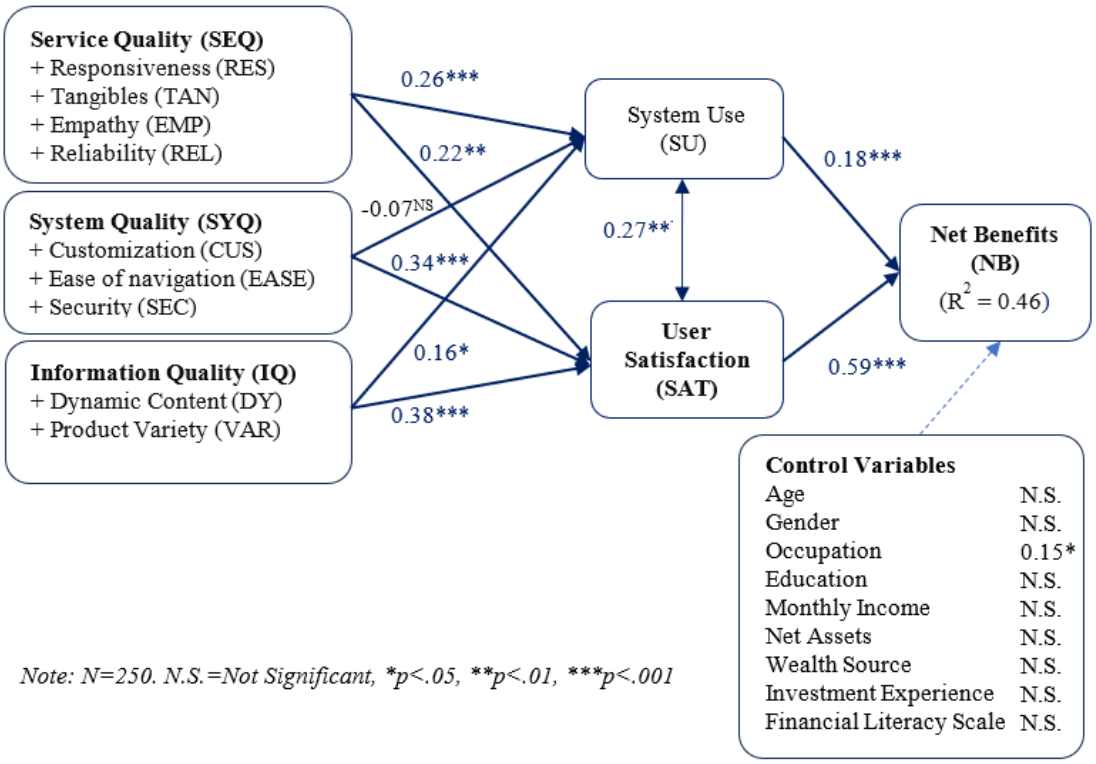


Figure 1. Research Model and Results

Table 1 – Moderation Analysis Results

Hypothesis		Path Coefficient (β)	P-value	Results
H7a	L*SEQ → SU	0.107 *	0.027	Supported
H7b	L*SEQ → SAT	-0.058	0.111	N.S.
H8a	BITs*SU → NB	-0.021	0.726	N.S.
H8b	BITs*SAT → NB	-0.007	0.897	N.S.

Table 2 - Multigroup (BITs) Analysis Results

Hypothesis		Original Group1 (PP&FF)	Original Group2 (II&AA)	t value Group1 (PP&FF)	t value Group2 (II&AA)	p-value Group1 (PP&FF)	p-value Group2 (II&AA)
H8a	SU -> NB	0.087	0.202***	0.955	3.686	0.339	0.000
H8b	SAT -> NB	0.581***	0.583***	6.487	11.410	0.000	0.000

*Note: N=250. N.S.=Not Significant, *p<.05, **p<.01, ***p<.001*

This study provides a number of contributions to theory and literature. First, this study expands the IS success model (Delone and McLean, 2003) within the context of DWPs by focusing on individual investors and introducing significant moderating factors: customer loyalty and BITs. Second, this study integrates customer loyalty and BITs as moderators to enrich the IS success model framework. Drawing

on the service-profit chain model (Heskett *et al*, 1994), it explores how customer loyalty influences the relationships between service quality and system use, as well as user satisfaction. Third, incorporating BITs (Pompian, 2008) allows the study to assess how psychological traits influence the links between system use and net benefits, and between user satisfaction and net benefits. Fourth, unlike traditional models that focus on adoption, this study investigates long-term success measures, specifically how user satisfaction and system use lead to net benefits like cost savings, portfolio diversification, enhanced returns, and improved financial literacy. This post-adoption perspective fills a gap in the DWP literature by examining the sustained value platforms provide. For practitioners, this study offers practical insights for DWPs to enhance system use and user satisfaction by focusing on service quality, system quality, and information quality, along with the influence of customer loyalty and BITs.

Keywords: Digital Wealth Management, Information Systems Success Model, Risk Tolerance of Investors, and Loyalty

REFERENCES

Boreiko, D. and Massarotti, F. (2020) How risk profiles of investors affect robo-advised portfolios. Frontiers in Artificial Intelligence 3: 60.

Deloitte (2020) The future of wealth management revisited. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/us-future-of-wealth-management-revisited-winter-2020.pdf.

DeLone, W. H. and McLean, E. R. (2003) The DeLone and McLean model of information systems success: a ten-year update. Journal of Management Information Systems 19(4): 9-30.

Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E. and Schlesinger, L. A. (1994) Putting the service-profit chain to work. Harvard Business Review 72(2): 164-174.

Pompian, M. M. (2008) Using behavioral investor types to build better relationships with your clients. Journal of Financial Planning 21(10): 64.

Keywords: Template, formats, instructions, length, conference publications

C3

Generative AI-Enabled Value Creation in Robo-Advisory: A Grounded Theory Multiple-Case Study

Qin Gao, Weihui Li, Yongqing Yang (Shandong Technology and Business University)

In the context of the digital economy, generative artificial intelligence (AI) has become a key technological driver for the transformation of robo-advisory services in the securities industry. To uncover the logic of value creation, this study adopts a multiple-case approach and applies grounded theory to analyze top securities firms ranked by total revenue in 2025.Through a three-stage coding process, three value creation patterns enabled by generative AI are identified: data-driven mechanisms, decision optimization, and service enhancement. Based on these findings, a theoretical model of generative AI-enabled value creation in robo-advisory is developed.This study enriches research on AI applications in fintech and provides theoretical support and practical insights for securities firms seeking to improve robo-advisory services and achieve value creation.

KEY WORDS

grounded theory, multiple-case study, robo-advisory, generative artificial intelligence, value creation

Generative AI-Enabled Value Creation in Robo-Advisory: A Grounded Theory Multiple-Case Study

Gao, Q.[a], Li, W. H.[b], and Yang, Y. Q.[b]
[a] School of International Business, Shandong Technology and Business University
E-mail: 37778774@qq.com
[b] School of Management Science and Engineering, Shandong Technology and Business University
E-mail: 1113753017@qq.com, yangyq@sdtbu.edu.cn

Abstract

In the context of the digital economy, generative artificial intelligence (AI) has become a key driver of transformation in robo-advisory services in the securities industry. To examine the logic of value creation, this study adopts a multiple-case approach and applies grounded theory to analyze top securities firms ranked by total revenue in 2025.Through a three-stage coding process, three value creation mechanisms enabled by generative AI are identified: data-driven, decision optimization, and service enhancement. Based on these findings, a theoretical model of generative AI-enabled value creation in robo-advisory is developed.This study contributes to research on AI applications in fintech and provides theoretical and practical insights for improving robo-advisory services and enhancing value creation in securities firms.

Keywords: grounded theory; multiple-case study; robo-advisory; generative artificial intelligence; value creation

Introduction

The rapid development of the digital economy and financial technology has positioned artificial intelligence (AI) as a key driver of transformation in wealth management. Robo-advisory, a core form of financial e-commerce service, aims to serve a broad user base, provide personalized investment matching, and deliver low-cost services. However, existing systems face structural challenges: limited data processing capabilities constrain analysis of unstructured financial data; algorithmic "black-boxes" reduce transparency and investor trust; and standardized decision processes fail to accommodate heterogeneous user preferences. Additionally, human advisors are costly and difficult to scale, while purely algorithm-driven services lack interaction, making it challenging to balance efficiency and personalization (Cardillo and Chiappini 2024).

Recent advances in generative AI demonstrate strong capabilities in natural language understanding, generation (Brown et al. 2020), and multimodal representation learning (Radford et al. 2021), offering potential solutions. Yet, research has not systematically explored how generative AI can address these challenges or its mechanisms for enabling value creation in digital financial services. This study investigates how generative AI enhances value creation in robo-advisory, focusing on its role in improving data processing, decision transparency, and service provision, thereby providing theoretical insights and practical guidance for technology-enabled digital service innovation.

Literature Review

Research on Value Creation

With the development of the resource-based view and value chain theory, scholarly understanding of value creation has shifted from single-factor drivers to the synergy of multiple elements. Amit and Zott (2001) highlight efficiency, complementarities, lock-in, and novelty as key sources of value creation, emphasizing multi-factor synergy. Building on this, Lepak et al. (2007) distinguish between value creation and value capture, proposing a multilevel framework encompassing individual, organizational, and network levels. Drawing on service-dominant logic, Vargo and Lusch (2004) argue that value is co-created through resource integration and service exchange among multiple actors.

The perspective on value creation has expanded from within firms to multi-actor collaboration. Value chain and stakeholder theories suggest that value originates from resource integration and relational coordination (Wang and Yin 2005; Wang 2010). In the digital economy, data and digital technologies have become key drivers, reshaping value creation through process reconfiguration and connecting multiple actors (Bharadwaj et al. 2013; Grover and Kohli 2012).

Research on the Application of Generative AI in Empowering Robo-Advisory

In the context of generative AI applications in robo-advisory, research has focused on its foundational logic and technical capabilities. Hong and Wang (2023) note that generative AI is shifting financial investment research from a "data-driven" to a "knowledge- and reasoning-driven" paradigm, leveraging LLMs' comprehension and reasoning capabilities to reshape information analysis and decision-making. Tahvildari (2025) further highlights its role in transforming robo-advisory from an "algorithm-driven" to a "cognition- and generation-driven" model, enhancing investment proposal generation and conversational services.

At the application level, studies emphasize technological implementation and service formats. Generative AI is applied in financial text processing, investment decision support, and client interaction (Jiang, Liu, & Meng, 2024; Adedoyin, Dogan, Cetinkaya, & Jiang, 2025). Overall, integrating generative AI with robo-advisory has become a key research direction in fintech, addressing both technical and practical aspects.

Research Gaps

Existing studies have advanced the understanding of value creation, the enabling role of digital technologies, and the basic applications of generative AI in robo-advisory. However, research at the intersection of these areas remains limited. Prior work on digital technology mainly focuses on traditional IT tools, while the value creation mechanisms of generative AI remain underexplored. In addition, existing studies on generative AI in robo-advisory emphasize technical implementation and system design, with limited attention to value creation processes and pathways. A systematic framework linking generative AI and value creation in robo-advisory is still lacking.

Research Design

Research Methodology

This study adopts a grounded theory approach with a multi-case empirical design to explore value creation logic in generative AI-enabled robo-advisory. Based on case data from leading securities firms, it uses systematic coding and case analysis to identify key AI capability elements and constructs a theoretical model of "AI capabilities–scenario adaptation–value creation."

The methodology is justified by three reasons. First, as an emerging topic without a mature framework, grounded theory enables theory building from empirical data (Strauss & Corbin, 1998). Second, multiple-case analysis allows cross-case comparison, improving robustness and

generalizability (Yin, 2018). Third, this approach captures the complex adaptation between AI capabilities and business scenarios, which is difficult for traditional quantitative methods (Mao & Zhang, 2008).

Case Selection

To ensure reliability and interpretive power, this study follows theoretical sampling principles (Glaser & Strauss, 1999) and selects cases based on industry representativeness, technological relevance, and data availability. The top ten securities firms by 2025 revenue—CITIC Securities, Guotai Haitong, Huatai, GF, China Galaxy, CICC, Shenwan Hongyuan, Guosen, CITIC Construction Investment, and China Merchants Securities—are chosen due to their strong industry influence and relevance to robo-advisory development. All selected firms disclose applications of generative AI in robo-advisory or related services, ensuring alignment with the focus on "AI capabilities–business application–value outcomes." Multi-source data from annual reports, official websites, white papers, and media reports support systematic coding and enhance research credibility. Overall, these firms are highly representative in technological application and digitalization, providing empirical and theoretical support for the generalizability of the findings.

Data Collection and Coding Process

Data Collection

This study relies primarily on secondary data and builds a multi-level data source system to enhance systematicity and interpretive power. Multi-source secondary data analysis is widely used in case studies and supports theory development (Su & Liu, 2013). Data were collected from major academic databases, including CNKI, Wanfang, Web of Science, and Scopus. Core data come from annual reports, technical white papers, and official disclosures of the ten sample securities firms, supplemented by industry reports, financial media, and professional platform materials. For data processing, sources were cross-checked to remove irrelevant information and duplicates, improving reliability and consistency. Data collection across the ten firms was balanced to reduce sample bias. The final dataset contains approximately 140,000 words, covering corporate reports, technical documents, industry reports, and media materials, providing a solid basis for coding and theory development.

Grounded Theory Coding Process

Open coding

Open coding is the initial step of grounded theory, focusing on the conceptualization and categorization of textual data. Through data decomposition and analysis, 51 initial concepts are identified. Based on semantic similarity and logical relationships, these concepts are grouped and further abstracted into 18 conceptual categories after merging and refinement. Examples of open coding are shown in Table 1.

Table 1. Example of open coding

Raw data	Initial Concepts	Concept Category
Guosen Securities' intelligent robo-advisory assistant covers high-frequency client inquiries such as business processing and trading rules, as well as employee support scenarios including stock analysis, fund diagnostics, and account evaluation, forming a multi-scenario intelligent robo-advisory solution.	24/7 Data Processing	Data Analysis Capability
Huatai Securities' "ZhangLe Smart Investment" relies on big data and AI algorithms to provide full-process	Data-Driven Asset	

asset allocation services for clients with different risk preferences and investment horizons.	Allocation	
Guotai Junan leverages large AI models to achieve massive data processing and automated production of investment advisory content.	Massive Data Processing	
Huatai Securities' "AI ZhangLe" is equipped with an AI-driven dynamic tracking engine that scans market-wide news, public sentiment, and announcements in real time, enabling rapid identification of trending themes and unexpected events.	Multi-Source Data Integration	
...	...	...

Axial coding

Following open coding, initial categories with explanatory power and emerging logical relationships are further integrated through axial coding. By clarifying causal linkages, hierarchical relations, and logical progression among categories, 18 conceptual categories are consolidated into three core categories: data-driven, decision optimization, and service enhancement. The results of axial coding are shown in Table 2.

Table 2. Axial coding results

Core category	Corresponding categories	Category description
Data-Driven	Data Analysis Capability	Data analysis capability involves processing and mining multi-source financial data to support robo-advisory.
	Customer Profile Modeling	Client profiling refers to constructing user labels based on investor behavior and risk preferences to identify personalized investment needs.
	Investment Opportunity Identification	Investment opportunity identification captures market trends from data and sentiment to support asset allocation and decisions.
	Multi-Source Data Integration	Multi-source data integration refers to integrating internal and external data to build a unified data foundation and improve analytical efficiency.
	Enhancing Decision Transparency	Decision transparency enhancement reduces algorithmic "black-box" effects via traceability and explainability, improving investor trust.
	Enhancing Analytical Precision	Analytical accuracy improvement refers to improving the accuracy of market judgment and risk assessment through data optimization and multi-model integration.
Decision Optimization	Intelligent Decision-Making Capability	Intelligent decision-making capability refers to leveraging generative AI to support investment analysis, asset allocation, and risk control.
	Robo-Advisory Plan Generation	Advisory plan generation creates personalized asset allocation plans based on investor characteristics.
	Automated Trade Execution	Automated trade execution refers to executing investment operations according to predefined strategies, improving efficiency and standardization.
	Robo-Advisory	Advisory strategy optimization refers to continuously

	Strategy Optimization	improving investment models through algorithm iteration and backtesting to enhance robustness.
	Decision Efficiency Enhancement	Decision efficiency improvement refers to reducing the time required for analysis, plan generation, and adjustment, thereby improving response speed.
	Risk Exposure Reduction	Risk exposure reduction lowers volatility and potential losses through risk control and monitoring.
Service Enhancement	Intelligent Service Capability	Intelligent service capability refers to integrating multi-scenario intelligent modules to provide full-cycle robo-advisory service support.
	Automated Report Generation	Automated report generation refers to using AIGC technologies to produce investment education content and analytical materials.
	Intelligent Customer Interaction	Intelligent customer interaction improves communication via large-model customer service and digital humans.
	Intelligent Post-Investment Tracking	Intelligent post-investment tracking refers to monitoring holdings, risks, and returns in real time and providing feedback.
	Customer Experience Improvement	Customer experience improvement refers to enhancing satisfaction and user stickiness through personalized and intelligent services.
	Service Scale Expansion	Service scale expansion reduces costs via digital services and expands coverage to long-tail clients.

Selective coding

Selective coding is a further theoretical abstraction based on axial coding, aimed at identifying the core category and integrating its relationships with other categories. This study defines "the value creation mechanism of generative AI-enabled robo-advisory" as the core category, forming a structured framework. Generative AI operates through three progressive pathways — data-driven, decision optimization, and service enhancement — forming a closed-loop process that ultimately enables value creation in robo-advisory services.

Theoretical saturation test

To ensure the completeness of the category system and theoretical saturation, this study conducts validation from two aspects: data coding and literature comparison. At the data coding level, continuous coding and repeated comparison reveal no new concepts, categories, or relationships, indicating that the existing categories sufficiently explain the core phenomenon. At the literature level, systematic review of related studies shows that the identified categories and relationships are well supported by existing theories, with no emergence of new theoretical dimensions.

Case Analysis and Model Development

Mechanisms of Value Creation in Generative AI-Enabled Robo-Advisory Services

Based on a multiple-case analysis and grounded theory coding of ten leading securities firms, this study finds that generative AI applications in robo-advisory show functional differences across scenarios, while remaining centered on core processes such as data processing, decision support, and service interaction. Based on capability configuration and value-creation pathways, three core value creation mechanisms are identified: data-driven, decision optimization, and service enhancement.

The data-driven mechanism is a foundational mechanism focused on improving data analytics capabilities to address inefficient use of unstructured data and low multi-source integration efficiency. Generative AI integrates heterogeneous financial data, analyzes user behavior, and supports client profiling. It reduces information asymmetry, improves analytical accuracy, and enhances decision transparency, mitigating the "black-box" problem and supporting downstream processes.

The decision optimization mechanism is the core transformation mechanism, addressing limitations in standardized decision-making and dynamic risk control. It improves the decision chain — covering strategy optimization, portfolio generation, and automated execution — through multi-model integration, RAG-based tracing, and reinforcement learning. This enhances efficiency, reduces latency, improves traceability, and strengthens risk control, supporting better performance, client retention, and asset growth.

The service enhancement mechanism is the terminal value realization mechanism, converting decision outputs into customer and business value through intelligent service capabilities. Generative AI enables automated reporting, interactive services, digital tools, and personalized content, improving responsiveness and interaction quality. It expands coverage to long-tail clients, enhances scalability, and supports customer experience and business growth.

Theoretical Model Construction of Generative AI-Enabled Robo-Advisory

Based on grounded theory coding of ten case firms, this study develops a theoretical framework for value creation in generative AI-enabled robo-advisory (Figure 1), following an AI capabilities → application scenarios → value creation logic. The AI capability layer includes data analysis, intelligent decision-making, and intelligent service capabilities, supporting the scenario layer. In the understanding stage, AI integrates multi-source data, builds client profiles, and identifies investment opportunities. In the decision-and-action stage, it supports advisory generation, automated trade execution, and strategy optimization. In the service stage, decisions are translated into client-perceived value through reporting, interaction, and post-investment follow-up. The progression of scenarios involves two transitions: understanding → decision-and-action reflects deeper application, while decision-and-action → service reflects value realization. Feedback from interaction, service performance, and decision execution refines data and models, forming a closed-loop system. At the value creation level, the framework enhances cognition, decision performance, and service value, enabling value co-creation for investors and firms. This model provides practical guidance for improving robo-advisory services.

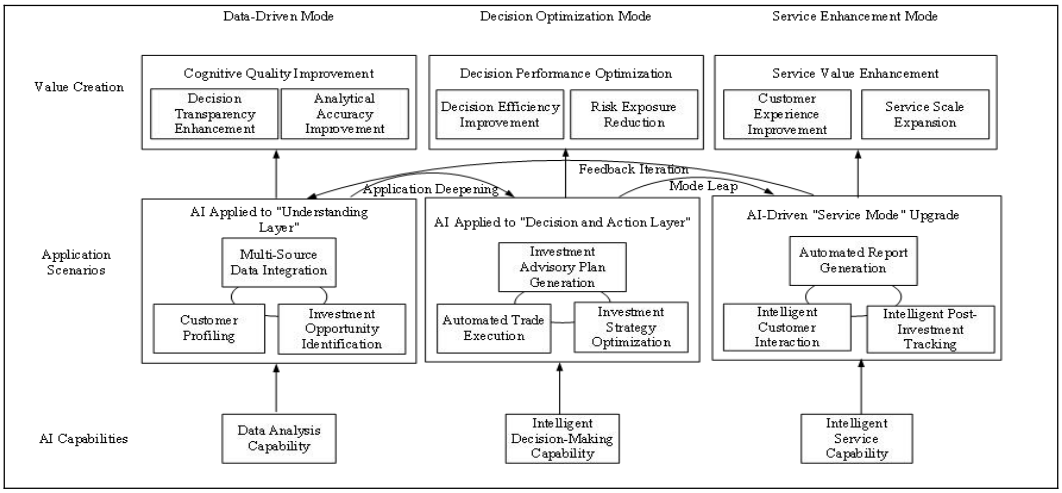


Figure 1. Theoretical Framework of Generative AI Empowering Robo-Advisory

The theoretical model of value creation in generative AI-enabled robo-advisory developed in this study is not only applicable to leading securities firms in China but also provides reference for various wealth management institutions with data and technological capabilities. Its layered logic of "AI capabilities – scenario adaptation – value creation" and closed-loop mechanism demonstrate strong theoretical generalizability and practical transferability.

Research Conclusions

Key Findings

Based on a multiple-case analysis, this study identifies three core value-creation mechanisms in generative AI-enabled robo-advisory: data-driven, decision optimization, and service enhancement, and develops a corresponding theoretical framework. Generative AI operates through a layered pathway: at the understanding layer, data integration, client profiling, and investment opportunity identification improve cognitive quality; at the decision-and-execution layer, portfolio generation, automated trading, and strategy optimization convert data into decision performance. At the service layer, automated reporting, intelligent interaction, and post-investment tracking translate decisions into customer value. Feedback from user behavior, decision execution, and service performance continuously refines data and cognitive models, forming a closed-loop mechanism. This mechanism enhances operational efficiency, risk management, and service quality, offering a replicable pathway for intelligent upgrading in digital service platforms.

Theoretical Contributions

This study advances theoretical research on generative AI in financial technology. While prior studies focus on either AI applications or individual robo-advisory functions, this research adopts a full business-chain perspective, systematically revealing the enabling mechanisms of generative AI across data cognition, decision execution, and service enhancement. It constructs a theoretical model of value creation in AI-powered robo-advisory, clarifying the formation logic and transmission of each value dimension. The findings provide a systematic framework for understanding value creation in scalable and inclusive digital financial services and offer a reference for future research on AI-enabled wealth management.

Practical Implications

For securities firms, generative AI implementation in robo-advisory should follow a progressive "cognition–decision–service" path, starting with data integration and client profiling, then advancing to decision automation and intelligent service, supported by a data feedback loop to iteratively optimize front-end models. Regulators should encourage innovation while ensuring AI decision transparency, data security, and investor protection. Small and medium-sized firms and digital platforms are advised to adopt lightweight deployment, focusing on low-cost scenarios such as intelligent customer service and personalized reports, gradually expanding to data governance and strategy optimization. Overall, generative AI can enhance inclusive finance and promote scalable, precise, and high-quality digital wealth management services.

Limitations and Future Research

This study has several limitations. First, although focusing on leading firms reflects mainstream industry practice, the generalizability of the findings to small and medium-sized brokerages, regional financial institutions, and cross-border wealth management scenarios requires further validation. Second, the study relies primarily on qualitative case analysis; future research could incorporate quantitative methods such as econometric models or structural equation modeling to improve robustness. Third, regulatory constraints and ethical risks are not fully addressed, leaving boundary conditions for further refinement. Future research may proceed in three directions. First, the case scope can be expanded to include small and medium-sized institutions, bank wealth management, and fund companies for cross-scenario comparison. Second, a quantitative indicator system can be

developed to empirically test the value-creation mechanisms of generative AI. Third, regulatory compliance and algorithmic ethics can be integrated into the analytical framework to better align with industry practice and long-term development.

References

Cardillo, G., and Chiappini, H. 2024. "Robo-Advisors: A Systematic Literature Review," Finance Research Letters (62), p. 105119.

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., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. 2020. "Language Models are Few-Shot Learners," in Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS), pp. 1877–1901.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. 2021. "Learning Transferable Visual Models from Natural Language Supervision," in Proceedings of the 38th International Conference on Machine Learning (ICML), pp. 8748–8763.

Amit, R., and Zott, C. 2001. "Value Creation in E-Business," Strategic Management Journal (22:6–7), pp. 493–520.

Lepak, D. P., Smith, K. G., and Taylor, M. S. 2007. "Value Creation and Value Capture: A Multilevel Perspective," Academy of Management Review (32:1), pp. 180–194.

Vargo, S. L., and Lusch, R. F. 2004. "Evolving to a New Dominant Logic for Marketing," Journal of Marketing (68:1), pp. 1–17.

Wang, H., and Yin, M. 2005. "A Study on the Element System of Value Creation under the Value Chain Model: A Comparison with Traditional Models in Value Evaluation," Guanli Shijie (Management World) (5), pp. 104–110, 143. (in Chinese)

Wang, S. 2010. "On the Fundamental Nature, Internal Mechanism, and Governance Implications of Value Creation: A Stakeholder Governance Perspective," Waiguo Jingji yu Guanli (Foreign Economics & Management) (32:8), pp. 10–17. (in Chinese)

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., and Venkatraman, N. 2013. "Digital Business Strategy: Toward a Next Generation of Insights," MIS Quarterly (37:2), pp. 471–482.

Grover, V., and Kohli, R. 2012. "Cocreating IT Value: New Capabilities and Metrics for Multifirm Environments," MIS Quarterly (36:1), pp. 225–232.

Hong, Y., and Wang, S. 2023. "Recent Advances in Artificial Intelligence and Their Impact on the Research Paradigm of Economics," Bulletin of Chinese Academy of Sciences (38:3), pp. 353–357. doi:10.16418/j.issn.1000-3045.20230303001.

Tahvildari, M. 2025. "Integrating Generative AI in Robo-Advisory: A Systematic Review of Opportunities, Challenges, and Strategic Solutions," Multidisciplinary Reviews.

Jiang, F., Liu, Y., and Meng, L. 2024. "Large Language Models, Text Sentiment, and Financial Markets," Management World (40:8), pp. 42–64. doi:10.19744/j.cnki.11-1235/f.2024.0090.

Adedoyin, F. F., Dogan, H., Cetinkaya, D., and Jiang, N. 2025. "Human-Centered AI in FinTech: A Conceptual Model and Strategic Research Agenda," Proceedings of the IEEE Conference on Artificial Intelligence (CAI), pp. 180–187.

Strauss, A. L., and Corbin, J. 1998. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (2nd ed.). Thousand Oaks, CA: Sage Publications.

Yin, R. K. 2018. Case Study Research and Applications: Design and Methods (6th ed.). Thousand Oaks, CA: Sage Publications.

Mao, J., and Zhang, X. 2008. "Standardization and Status Evaluation of Case Study Research: A Review of the China Enterprise Management Case Forum (2007)," Management World (4), pp. 115–121. doi:10.19744/j.cnki.11-1235/f.2008.04.012.

Glaser, B. G., and Strauss, A. L. 1999. The Discovery of Grounded Theory: Strategies for Qualitative Research. London: Routledge.

Su, J., and Liu, J. 2013. "Research on the Reliability of Secondary Data from the Perspective of Case Study Standardization," Chinese Journal of Management (10:10), pp. 1405–1409, 1418.

KITS·ICEC·KIISS
2026 통합 국제학술대회

C4 Human Interaction & Future Intelligent Environments

Session Chair : Wenlong Zhu
(Qingdao Univ. of Technology)



C4

How Does Human-AI Fit Shape Consumers' Passive Decision-Making?

Wenlong Zhu, Lei Liu (School of Business, Qingdao University of Technology, Qingdao, Shandong, China)

Artificial intelligence (AI) has reshaped decision-making patterns of e-commerce consumers, yet the formation mechanism of passive decision-making driven by human-AI fit remains underexplored. This study extracts 6 core variables via text mining of user comments from Xiaohongshu and Douyin, and constructs a three-stage cognitive-affective-passive decision-making model. Structural equation modeling is adopted for empirical analysis on 453 valid questionnaires from AI-enabled e-commerce users. Results showed that algorithm adaptation exerts significant positive effects on emotional intimacy, emotional security, and emotional resonance, serving as a core driver of users' emotional formation. Emotional intimacy, emotional security and emotional resonance each have significant positive effects on recommendation adoption behavior and judgment concession behavior. This study enriches AI e-commerce consumer behavior research and provides practical strategies for enterprises to optimize algorithms and build emotion-driven human-AI interaction systems.

KEY WORDS

Artificial Intelligence, Passive Decision-Making, Human-AI Fit, Cognitive-Affective-Behavior Theory

How Does Human–AI Fit Shape Consumers' Passive Decision-Making?

Zhu, W. L.[a] and Liu, L.[b]
[a] Business School, Qingdao University of Technology
E-mail: zhuwenlong1985@163.com
[b] Business School, Qingdao University of Technology
E-mail:15610060855@163.com

Abstract

Artificial intelligence (AI) has reshaped decision-making patterns of e-commerce consumers, yet the formation mechanism of passive decision-making driven by human-AI fit remains underexplored. This study extracts 6 core variables via text mining of user comments from Xiaohongshu and Douyin, and constructs a three-stage cognitive-affective-passive decision-making model. Structural equation modeling is adopted for empirical analysis on 453 valid questionnaires from AI-enabled e-commerce users. Results showed that algorithm adaptation exerts significant positive effects on emotional intimacy, emotional security, and emotional resonance, serving as a core driver of users' emotional formation. Emotional intimacy, emotional security and emotional resonance each have significant positive effects on recommendation adoption behavior and judgment concession behavior. This study enriches AI e-commerce consumer behavior research and provides practical strategies for enterprises to optimize algorithms and build emotion-driven human-AI interaction systems.

Keywords: Artificial Intelligence, Passive Decision-Making, Human–AI Fit, Cognitive-Affective-Behavior Theory

Introduction

China's online shopping market has reigned as the world's largest for 13 consecutive years, boasting a nationwide online retail sales volume of 15.9722 trillion yuan in 2025. The number of online shoppers in China has reached 976 million, making up 86.9% of the total internet user base. Amid the continuous expansion of this market, e-commerce competition has intensified dramatically, and artificial intelligence (AI) has emerged as a pivotal force reshaping the industry ecosystem by being widely integrated into every stage of the online shopping journey. From personalized product recommendations and AI-powered customer service to shopping assistants built on natural language processing, AI permeates all facets of e-commerce operations. By synthesizing and analyzing multi-dimensional user behavioral data—including browsing history, search records, and purchase preferences—AI enables precise human-machine matching, which in turn drives the dynamic alignment of services with users' actual needs and interaction habits. This technological integration not only reshapes consumers' decision-making processes and behavioral patterns but also significantly boosts service efficiency and elevates the overall user experience (Alghaswyneh, 2025).

Yet, as algorithm-driven interaction models gain increasing prevalence, a novel consumer behavior has come to the fore: passive decision-making. Characterized by a gradual erosion of users' motivation for independent decision-making and a growing reliance on algorithmic recommendations, this behavior arises from the sustained guidance of AI's precise matching mechanisms (Dul, 2024). Current academic research on AI-empowered e-commerce largely centers on the positive value propositions of the technology, such as enhanced operational efficiency and optimized user experience. Far less attention has been paid to the latent issue of passive decision-making stemming from human-machine adaptation, and the underlying logic of how this behavior forms remains

underexplored and unclear. In light of this research gap, the core research question of this study is formulated as follows: How does AI-driven human-machine fit induce consumers' passive decision-making behaviors in the context of AI-empowered e-commerce?

To address this question, this study first extracts the core variables that influence consumers' passive decision-making by analyzing authentic user comments and human-machine interaction feedback related to AI recommendations and intelligent customer service from leading platforms including Xiaohongshu and Douyin, on the basis of which a theoretical model is constructed. Structural equation modeling (SEM) is then employed to empirically examine the significant causal paths through which human-machine fit exerts an impact on consumers' passive decision-making. The findings of this research are expected to enrich the theoretical system of consumer behavior studies in AI-empowered scenarios. They also aim to provide e-commerce enterprises with both theoretical foundations and practical strategies to refine recommendation algorithm logic, enhance the quality of user interactions, and implement precision tiered operations.

This study makes two key theoretical contributions. First, it shifts from the positive bias in prior AI e-commerce research to focus on passive decision-making derived from human – AI fit, enriching research perspectives and filling gaps in passive decision behavior. Second, based on the cognition–affect–behavior framework, this study extracts core variables and constructs a three-stage model to explain how human – AI fit influences passive decision-making, deepening understanding of its formation mechanism.

Theoretical Foundation

This study is grounded in cognitive-affective-behavior (C-A-B) theory, a classic and widely adopted framework for explaining attitude formation and behavioral prediction. The theory posits that individual behavior follows a sequential, causal, and logically consistent transmission chain: cognitive evaluation → affective arousal → behavioral response (Dul, 2024; Liu et al., 2021). Cognition refers to individuals' rational perceptions, judgments, and sense-making of external stimuli or objects. Affect represents the emotional responses, feelings, and psychological states generated on the basis of cognitive processing. Behavior constitutes the overt, observable reactions driven by the combined effects of cognition and affect. This hierarchical chain highlights that behavior is not formed randomly but emerges through a progressive conversion from rational perception to emotional experience, and finally to action.

In the context of AI-enabled e-commerce, the C-A-B framework provides a clear and consistent theoretical logic for analyzing how users' interactions with AI systems drive their final decision-making responses. Specifically, users first form cognitive perceptions of the AI system's performance and matching quality. On that basis, they generate corresponding affective reactions and psychological states. These cognitive and affective experiences together further shape their final decision-making behaviors.

By following this stable "perception – emotion – behavior" transmission path, the C-A-B theory ensures that the research model maintains rigorous theoretical causality rather than loose variable correlation. It also offers a systematic theoretical lens for revealing the formation mechanism of consumers' passive decision-making in human-AI interaction scenarios, thereby enhancing the overall theoretical validity and explanatory power of this study.

Variable Extraction and Research Hypotheses

Text Mining and Core Variable Extraction

We collected user comments regarding AI recommendations and intelligent customer service from Xiaohongshu and Douyin. A total of 11,318 comments were crawled via Python in compliance with platform protocols and data security regulations, covering only public and non-sensitive content. After preprocessing, 9,027 valid comments were retained. Jieba segmentation, stop-word filtering, and synonym merging were adopted for text analysis. Following the thematic analysis procedure, two

researchers coded high-frequency keywords independently, with an inter-coder consistency coefficient above 90%(Kacprzak & Hensel ,2023). Accordingly, core variables consistent with the cognitive-affective-behavior framework were systematically extracted.

Variable Extraction and Delineation

Drawing on the semantic connotations of high-frequency keywords, results of thematic clustering and established theoretical foundations, this study systematically extracted the core variables within the cognitive-affective-behavioral three-stage framework. By means of quantitative and qualitative analyses of text mining, the screening, sorting and definition of variables were completed. The core variables finally extracted are presented as follows, which lay a solid variable foundation for the subsequent model construction and empirical analysis.

1. Cognitive stage refers to human-ai fit perception:
Algorithmic understanding: Consumers' perception of the algorithm's ability to understand their needs, emotions and shopping goals;
Algorithmic alignment: The matching degree between the AI recommendation system and consumers' consumption preferences, values and shopping goals;
Algorithmic adaptation: The flexibility of the algorithm to adjust recommendation content according to consumers' behavioral feedback, preference changes and real-time needs.
2. Affective stage refers to AI-driven emotional connection:
Emotional intimacy: Subjective psychological experience of "familiarity and close emotional connection" generated by AI's timely, personalized and anthropomorphic responses;
Emotional security: Psychological stability, emotional relaxation and decision-making confidence obtained by AI's clarity, professionalism and support;
Emotional resonance: Irrational but stable emotional identification and dependence formed by consumers' familiarity with the algorithm operation mechanism and cumulative experience convenience.
3. Behavioral stage refers to passive decision-making performance:
Recommendation adoption behavior: Passive behavior where consumers directly accept AI recommendations and take them as decision options or final choices;
Judgement concession behavior: Passive decision-making where consumers entrust commodity evaluation and information screening to AI and abandon independent judgment.

Research Hypotheses

The Impact of Cognitive Stage on Affective Stage

Algorithmic understanding is consumers' perception of AI's ability to identify their deep-seated needs, emotional tendencies and shopping goals. This perception makes consumers feel valued and recognized, enhancing their psychological identity with AI to foster emotional intimacy, reducing shopping uncertainty to improve emotional security, and aligning cognitive expectations with AI's performance to drive emotional resonance (Cuesta-Valino et al., 2025). In view of this, this study proposes the following hypotheses:

H1: Algorithmic understanding has a significant positive impact on emotional intimacy.
H2: Algorithmic understanding has a significant positive impact on emotional security.
H3: Algorithmic understanding has a significant positive impact on emotional resonance.

Algorithmic alignment reflects the matching degree between AI's operational logic and consumers' consumption preferences, values and usage habits. Such consistency reduces human-machine interaction costs and psychological estrangement to boost emotional intimacy, forms a psychological buffer against decision-making anxiety to improve emotional security, and brings a sense of familiarity to narrow the human-AI psychological distance and trigger emotional resonance (Gerlich, 2023). In view of this, this study proposes the following hypotheses:

H4: Algorithmic alignment has a significant positive impact on emotional intimacy.
H5: Algorithmic alignment has a significant positive impact on emotional security.
H6: Algorithmic alignment has a significant positive impact on emotional resonance.

Algorithmic adaptation, the core of AI's co-evolution with consumers, refers to its dynamic adjustment of services based on real-time consumer feedback and changing demands. Unlike static cognitive fit variables, its dynamic optimization deepens consumers' emotional investment to form emotional intimacy (Nagy & Hajdú, 2021), its stable performance buffers trust decline to underpin emotional security, and its anticipation of changing needs sustains emotional interaction to deepen emotional resonance, acting as a key driver of multidimensional emotional bonds. In view of this, this study proposes the following hypotheses:

H7: Algorithmic adaptation has a significant positive impact on emotional intimacy.

H8: Algorithmic adaptation has a significant positive impact on emotional security.

H9: Algorithmic adaptation has a significant positive impact on emotional resonance.

The Impact of Affective Stage on Behavioral Stage

Emotional intimacy is consumers' subjective experience of close emotional connection with AI derived from its anthropomorphic and personalized responses. This emotional bond strengthens consumers' dependence on AI, reduces resistance to AI recommendations to promote recommendation adoption behavior, and increases willingness to entrust commodity evaluation and information screening to AI to drive judgement concession behavior (Xiao & Benbasat, 2018). In view of this, this study proposes the following hypotheses:

H10: Emotional intimacy has a significant positive impact on recommendation adoption behavior.

H11: Emotional intimacy has a significant positive impact on judgement concession behavior.

Emotional security is the psychological stability and decision-making confidence obtained by consumers from AI's professional and reliable performance. As a key psychological guarantee for passive decision-making, it relieves decision-making anxiety to improve the willingness of recommendation adoption (Liu et al, 2021), and enhances trust in AI's information processing ability, making consumers abandon independent judgment and transfer decision-making power to AI, thus facilitating judgement concession behavior. In view of this, this study proposes the following hypotheses:

H12: Emotional security has a significant positive impact on recommendation adoption behavior.

H13: Emotional security has a significant positive impact on judgement concession behavior.

Emotional resonance is the stable emotional identification formed by consumers from long-term AI usage experience, serving as the core driver of passive decision-making. It eliminates human-machine interaction barriers to greatly enhance consumers' willingness to accept AI recommendations, and deepens trust in AI's decision-making tools to improve the willingness to transfer judgment power, even forming stable decision-making dependence on AI (Lim & Kim, 2021). In view of this, this study proposes the following hypotheses:

H14: Emotional resonance has a significant positive impact on recommendation adoption behavior.

H15: Emotional resonance has a significant positive impact on judgement concession behavior.

Based on the above hypotheses, this study constructed the theoretical model depicted in Figure 1.

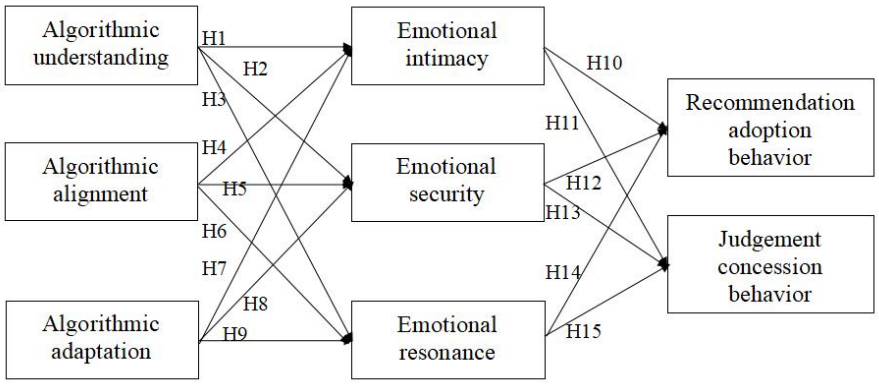


Figure 1. Theoretical model

Research Methodology

Measurement Scale

Based on the theoretical model and existing high-quality SSCI journal items, we designed a measurement scale with 8 variables and 30 items, using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). All items are referenced to mature scales in the existing literature (Wen & Li, 2025), ensuring the reliability and validity of the scale.

Data Collection

We selected active consumers who regularly utilize AI-enabled features on major e-commerce platforms including Tmall and JD as the research subjects. Questionnaires were distributed via the Credamo platform from December 5 to December 30, 2025, with a total of 567 questionnaires issued. After rigorously screening out invalid responses—such as those with extensive repetitive answers—a total of 453 valid questionnaires was retained for subsequent analysis, yielding an effective response rate that meets the requirements of empirical research. The demographic profile of the valid sample is as follows: 54.97% of respondents were female and 62.69% aged between 18 and 30 years old; 72.41% held a bachelor's degree or lower educational qualification. In terms of online shopping behavior, 49.23% of participants shopped online more than three times a week, and 83.44% reported frequently or almost always using AI functions during their shopping process. Overall, the sample characteristics are highly consistent with the typical user profile of AI-empowered e-commerce platforms, ensuring the representativeness and validity of the research data.

Empirical Results

Reliability and Validity Test

Exploratory factor analysis was conducted via SPSS 27.0, yielding a KMO value of 0.914, which confirms the suitability of the data for factor analysis. Item AF3 was eliminated due to a factor loading below 0.7, while all remaining items achieved factor loadings in excess of 0.7, meeting the established academic threshold (Westland, 2010). For reliability assessment, both Cronbach's α and composite reliability (CR) values for all variables exceeded 0.7, demonstrating strong internal consistency of the measurement scale. Validity tests further verified the scale's psychometric quality: the average variance extracted (AVE) for each variable was above 0.5, confirming satisfactory convergent validity; additionally, the square root of each variable's AVE was greater than its correlation coefficients with other constructs, which attested to good discriminant validity across all dimensions. Overall, these results validate the high reliability and validity of the research scale (Table 1).

Table 1. Scale Reliability and Validity Analysis Results (Partial)

Variable	Cronbach's α	CR	AVE	Factor Load Range
Algorithmic Understanding	0.90	0.88	0.70	0.82-0.86
Algorithmic Alignment	0.91	0.83	0.71	0.84-0.84
Algorithmic Adaptation	0.95	0.89	0.67	0.74-0.85
Emotional Intimacy	0.95	0.95	0.79	0.84-0.93
Emotional Resonance	0.87	0.84	0.56	0.72-0.78
Recommendation Adoption	0.86	0.86	0.67	0.74-0.86
Judgement Concession	0.82	0.87	0.58	0.71-0.81

SEM Hypothesis Test

Model fit indices confirmed a good fit of the theoretical model to the research data ($\chi^2/df=2.55<3$, RMSEA=0.074<0.08, NFI=0.95>0.9, CFI=0.97>0.9, GFI=0.91>0.9). Hypothesis testing results revealed that 12 out of the 15 proposed hypotheses were supported, with three hypotheses (H1, H2, H4) not being validated (see Figure 2).

From the cognitive to the affective stage: Algorithmic understanding exerted a significant positive effect only on emotional resonance (H3 supported). Algorithmic alignment had a significant positive impact on both emotional security and emotional resonance (H5, H6 supported). Notably, algorithmic adaptation demonstrated a significant positive influence on all three affective variables (H7-H9 supported), emerging as the core cognitive construct driving the formation of consumers' emotional bonds with AI systems.

From the affective to the behavioral stage: Emotional intimacy, emotional security, and emotional resonance all exerted significant positive effects on both recommendation adoption and judgement concession behaviors (H10-H15 all supported). These results empirically verify the critical mediating role of the affective stage in the transmission mechanism from human-AI cognitive fit to consumers' passive decision-making behaviors.

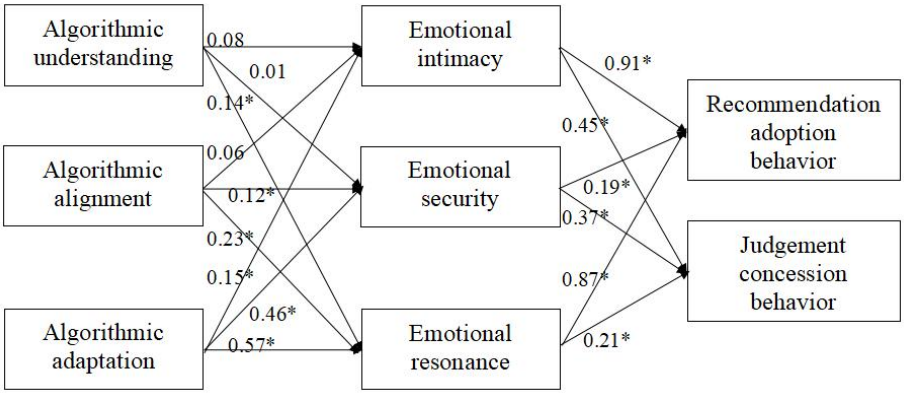


Figure 2. Hypothesis testing results

Conclusions and Implications

Conclusions

First, human–AI fit has a significant positive impact on consumers' emotions. Specifically, algorithmic understanding positively affects emotional resonance; algorithmic alignment enhances emotional security and emotional resonance; and algorithmic adaptation significantly promotes emotional intimacy, emotional security, and emotional resonance. These results are consistent with prior studies (Nagy & Hajdú, 2021). However, neither algorithmic understanding nor alignment significantly improves emotional intimacy, which differs from existing research (Gerlich, 2023; Cuesta-Valino et al., 2025). Algorithmic understanding and algorithmic alignment, as forms of relatively static human-AI cognitive fit, merely trigger partial affective responses in consumers, while dynamic algorithmic adaptation can effectively deepen emotional bonds between humans and AI by accommodating consumers' evolving needs. The key advantage of algorithmic adaptation lies in its long-term, dynamic matching with consumers, rather than superficial need identification or behavioral alignment.

Second, consumers' emotions significantly promote passive decision-making. Emotional intimacy, emotional security, and emotional resonance all positively affect recommendation adoption and judgement concession, supporting previous findings (Lim & Kim, 2021; Liu et al, 2021; Xiao & Benbasat, 2018). Emotion acts as a critical mediator between cognitive perceptions of human–AI fit and passive decision-making. Emotional intimacy encourages psychological acceptance of AI; emotional security reduces resistance and perceived risk; emotional resonance strengthens trust and

interaction stickiness. These findings confirm the applicability of the C-A-B theory in AI-powered e-commerce and demonstrate that passive decision-making is a rational choice based on emotional connection, rather than simple cognitive inertia.

Implications

Theoretical Implications

First, this study clearly classifies consumers' passive decision-making behaviors into recommendation adoption and judgement concession, accurately deconstructing their manifestations and characteristics. It addresses the research gap in existing studies, which mostly explore such behaviors at a macro level without specific decomposition, leading to a lack of pertinence and depth. Through text mining and empirical analysis, this study clarifies the connotation and boundaries of the two behavioral types, enriches the knowledge system of passive decision-making in AI-empowered e-commerce, and provides a clear theoretical reference for subsequent research on behavior measurement, variable definition and mechanism analysis.

Second, the constructed "cognition-affect-passive decision-making" three-stage theoretical model, based on the C-A-B theory, systematically reveals the sufficient conditions and transmission paths of passive decision-making formation. It clarifies how cognitive variables (algorithmic understanding, alignment, and adaptation) influence the two passive decision-making behaviors through the mediation of emotional intimacy, emotional security, and emotional resonance. This model breaks the limitations of a single research method and provides a reference framework for subsequent studies to systematically analyze the impact of human-machine interaction on consumers' decision-making from the full cognition-affect-behavior chain.

Practical Implications

First, enterprises should establish a dynamic cognitive fit system centered on algorithmic adaptation and supported by algorithmic alignment. They should abandon static matching logic and build a three-tier structure based on algorithmic understanding, alignment, and adaptation. Priority should be given to enhancing real-time tracking of user preference changes and dynamic optimization of recommendation strategies using multi-dimensional behavioral data. Meanwhile, deepening the precise matching value of algorithmic alignment based on user portraits can lay a solid cognitive foundation for emotional connection.

Second, enterprises should build an emotion - driven human - AI interaction system to nurture consumers' passive decision-making via multi-dimensional emotional bonds. As the critical mediator between cognitive fit and passive behavior, the affective stage demands a shift from pure functional design to a full - process emotional experience centered on recommendation adoption and judgment concession. Firms should deliver warm, responsive, and trustworthy interactions through intelligent customer service and recommendation assistants. For emotional intimacy, personalized memory and tailored responses can strengthen attachment and willingness to adopt suggestions. For emotional security, transparent recommendation logic and data rules enhance perceived professionalism and reduce resistance. For emotional resonance, algorithms should capture dynamic preferences to deepen emotional connection and reliance. Together, these three emotional dimensions effectively link cognitive fit with passive decision - making, significantly boosting recommendation acceptance and judgment concession.

Potential Negative Impacts and Ethical Considerations

While human – AI fit effectively fosters emotional connection and passive decision-making, over-reliance on algorithm-driven recommendations may entail potential negative consequences and ethical challenges. Excessive algorithmic adaptation and emotional resonance could gradually erode consumers' independent thinking and rational judgment, resulting in long-term decision inertia and cognitive inertia. Ethically, opaque algorithmic logic, excessive collection of user behavioral data, and one-sided consumption guidance may trigger issues including privacy violations, information

cocoons, and unfair consumption manipulation. Furthermore, intensive human–AI emotional bonding may lower consumers' risk awareness, rendering them more vulnerable to manipulative consumption behaviors. Therefore, enterprises should balance user experience with ethical principles, enhance algorithm transparency and data security, safeguard consumers' rights to independent decision-making, and avoid excessive guidance that leads to irrational or excessive passive reliance.

Research Limitations and Future Directions

There are two limitations in this study. First, although this study focuses on Chinese e-commerce platforms, the core cognitive–affective–behavior mechanism is highly generalizable to global AI e-commerce scenarios. The theoretical logic of human–AI fit driving emotions and passive decision-making aligns with universal consumer psychology. Cultural values and platform environments may moderate effect strength but not the core mechanism.Future cross-country comparative studies can be conducted to verify the cross-cultural applicability of the model and conclusions. Second, this study focuses on general e-commerce AI scenarios. Subsequent research may examine the heterogeneous effects of different AI forms (e.g., AI digital humans, intelligent voice assistants) and vertical e-commerce contexts, so as to further enrich the application boundaries of the research conclusions.

References

Alghaswyneh, O. F. 2025. "Exploring the Impact of AI-Driven Personalization on Consumer Engagement in Digital Marketing," *Management* (29:1), pp. 117-144.

Cuesta-Valino, P., Kazakov, S., Duran-Alamo, P., et al. 2025. "The Role of AI as an Unconventional Salesperson in Consumer Buying Decisions, Satisfaction and Happiness," *European Research on Management and Business Economics* (31:2).

Dul, J. 2024. "A Different Causal Perspective with Necessary Condition Analysis," *Journal of Business Research* (177), pp. 114618.

Fan, Y., and Liu, X. 2022. "Exploring the Role of AI Algorithmic Agents: The Impact of Algorithmic Decision Autonomy on Consumer Purchase Decisions," *Frontiers in Psychology* (13), pp. 1009173.

Gerlich, M. 2023. "The Power of Personal Connections in Micro-Influencer Marketing: A Study on Consumer Behaviour and the Impact of Micro-Influencers," *Transnational Marketing Journal* (11:1), pp. 131-152.

Kacprzak, A., and Hensel, P. 2023. "Exploring Online Customer Experience: A Systematic Literature Review and Research Agenda," International Journal of Consumer Studies (47:6), pp. 2583-2608.

Lim, S., and Kim, M. 2025. "AI-Powered Personalized Recommendations and Pricing: Moderating Effects of Ethical AI and Consumer Empowerment," *International Journal of Hospitality Management* (130).

Liu, B., Xu, Y., Yang, Y., et al. 2021. "How Public Cognition Influences Public Acceptance of CCUS in China: Based on the ABC (Affect, Behavior, and Cognition) Model of Attitudes," *Energy Policy* (156).

Nagy, S., and Hajdú, N. 2021. "Consumer Acceptance of the Use of Artificial Intelligence in Online Shopping: Evidence from Hungary," *Amfiteatru Economic* (23:56), pp. 155-173.

Wen, J., and Li, X. 2025. "AI Digital Human Responsiveness and Consumer Purchase Intention: The Mediating Role of Trust," *Journal of Theoretical and Applied Electronic Commerce Research* (20:3), pp. 246.

Westland, J. C. 2010. "Lower Bounds on Sample Size in Structural Equation Modeling," *Electronic Commerce Research and Applications* (9:6), pp. 476-487.

Xiao, B., and Benbasat, I. 2018. "An Empirical Examination of the Influence of Biased Personalized Product Recommendations on Consumers' Decision Making Outcomes," *Decision Support Systems* (110), pp. 46-57.

C4

BUILDING TRUST AND TIES: WHY ELDERLY PATIENTS CHOOSE ONLINE HEALTH COMMUNITIES

Jinxin Yang (Hefei Technology College), Xinyi Xie (King's College of London)

This study examines factors affecting elderly patients' intention to use Online Healthcare Communities (OHCs), emphasizing social interaction ties, shared value, and trust. Data from 307 elderly OHC users in China were analyzed using Partial Least Squares (PLS-SEM). Findings reveal that social identity, perceived critical mass, and information/emotional interactions strengthen social interaction ties. These ties positively influence shared value and trust, which subsequently enhance adoption intention. OHC managers should promote social interaction and address both informational and emotional needs to increase adoption. This research integrates social capital and motivation theories to offer a framework for improving digital healthcare engagement among the elderly.

KEY WORDS

Online Healthcare Community, Theory of Motivation, Social Capital Theory, Partial Least Squares

BUILDING TRUST AND TIES: WHY ELDERLY PATIENTS CHOOSE ONLINE HEALTH COMMUNITIES

Jinxin Yang [a] *and Xinyi Xie* [b]
[a] *School of Economics and Trade, Hefei Technology College,*
Email: 461004529@qq.com
[b] *Institute of Psychiatry, Psychology & Neuroscience, King's College of London, China,*
Email: liona911@icloud.com

Abstract

This study examines factors affecting elderly patients' intention to use Online Healthcare Communities (OHCs), emphasizing social interaction ties, shared value, and trust. Data from 307 elderly OHC users in China were analyzed using Partial Least Squares (PLS-SEM). Findings reveal that social identity, perceived critical mass, and information/emotional interactions strengthen social interaction ties. These ties positively influence shared value and trust, which subsequently enhance adoption intention. OHC managers should promote social interaction and address both informational and emotional needs to increase adoption. This research integrates social capital and motivation theories to offer a framework for improving digital healthcare engagement among the elderly.

Keywords: Online Healthcare Community, Theory of Motivation, Social Capital Theory, Partial Least Squares

Introduction

OHCs' usage increased among elderly patients during COVID-19, providing a platform for health information sharing and support to improve well-being(Boon-itt, 2019). Elderly patients' OHC interactions and information exchange enhance social capital, affecting their continued usage intention(Wu and Lu, 2016; Zhou, 2020).

Due to physical limitations hindering traditional healthcare access, elderly patients use OHCs for guidance, support, and health information in a safe, convenient environment(Zhou, 2023). These communities facilitate patient-doctor and patient-patient exchanges, meeting emotional and peer support needs beyond medical aspects(Atanasova et al., 2018). This study examines elderly patients' OHC adoption factors, focusing on shared value, trust, and social capital.

Previous research highlights perceived social capital and trust as key determinants of virtual health community participation(Meng et al., 2022). For isolated elderly users, OHCs build community and

trust in online healthcare (Ebardo and Suarez, 2021). Social capital developed through information/emotion interaction and trust motivates OHC participation, enhancing care quality and user experience(Connolly et al., 2023; Ebardo and Suarez., 2021).

This study proposes a model to examine how elderly patients' OHC adoption intentions are influenced by social interactions enhancing shared value, trust, and perceived social capital(Boon-itt, 2019). Findings aim to contribute to academic understanding and healthcare practitioners' efforts to enhance elderly digital healthcare engagement.

Theoretical Background and Hypotheses Development

Drivers of Social Interaction Ties

Based on Motivation Theory, drivers include extrinsic (social identity, information interaction) and intrinsic (perceived critical mass, emotional interaction) factors. Social identity fosters belonging and experience sharing; information interaction builds credibility through reciprocal health knowledge exchange, both strengthening social ties.

H1: Social identity positively impacts social interaction ties.

H3: Information interaction positively impacts social interaction ties.

Intrinsic drivers: Perceived critical mass enhances platform confidence via bandwagon effects, encouraging participation. Emotional interaction reduces isolation through shared stories, building trust and long-term engagement.

H2: Perceived critical mass positively impacts social interaction ties.

H4: Emotional interaction positively impacts social interaction ties.

Social Capital and Adoption Intention

Social capital enables OHC relationship building and resource exchange via shared health content, aligning with social support frameworks to foster capital accumulation and network expansion.

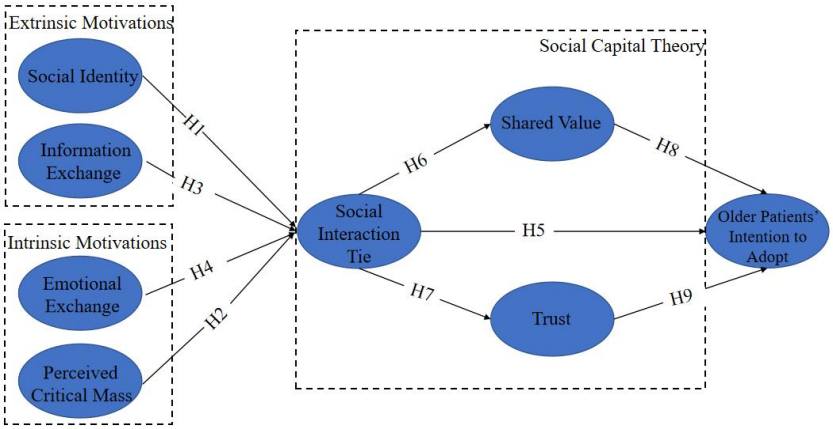


Figure1. Structure of the study

H5: Social interaction ties have a significant positive impact on elderly patients' intention to adopt OHCs.

H6: Social interaction ties have a significant positive impact on shared value.

H7: Social interaction ties have a significant positive impact on trust.

In OHCs, shared values and trust—cognitive social capital aspects—encourage platform use and communication. Trust motivates e-health service selection, leading to hypotheses:

H8: Shared value has a significant positive effect on elderly patients′ intention to adopt OHCs.

H9: Trust has a significant positive impact on elderly patients′ intention to adopt OHCs.

Methodology

Selection of the Cases and Data Collection

The study surveyed elderly OHC users in Hefei, China, using Thomas and Brubaker's method (Tompson*et al*., 1995) to assess adoption intention via anonymous online questionnaires distributed through WeChat groups from five local hospitals. From January to March 2023, 351 patients were surveyed; 44 were excluded due to incomplete data, resulting in 307 valid questionnaires (87.46% response rate). Among respondents, 159 (51.79%) were male and 148 (48.21%) female, aged 60-82 (average 73.3). Education levels: junior college (27.36%) and undergraduate (38.11%) were most common. Occupations: civilian staff (14.98%) was highest. OHC usage: 38.11% used 1-2 hours per session. Most were from second-tier or below cities (52.44%).

Variables and Measurement



Figure 2. Social interaction tie in OHCs: A case study of elderly specialists

Scales from validated literature were adapted for elderly patients by managers and doctoral students. The study includes 8 constructs: one with 6 items and seven with 3 each, totaling 27 items. All variables use a 7-point scale (1=Completely disagree, 7=Completely agree).

Platform-based medical services offer flexible, diverse options for elderly patients, especially those with chronic conditions, removing time and location barriers. Patients can review doctors' history, such as patient numbers, follow-up rates, and recommendation scores from platform data, which supports recommendations. Platforms note that doctors with more patients often have higher follow-up

rates, reinforcing key perceptions. Emotional interaction is enabled via doctors' communication accounts, allowing patients to send gifts post-consultation to express gratitude, per Figure 2.

Data Analysis and Results

Common Method Deviation

To avoid common method deviation, Harman's Single-Factor Test was employed (8 dimensions, 27 items) (Hoffman and Novak, 1996). Using eigenvalue >1 as criterion, the first factor explained 37.01% variance (below 40% standard) (Hair*et al*., 1998), indicating no serious common method deviation.

Framework quality was assessed via goodness of fit (GOF) using (Tenenhaus et al., 2005) mathematical model:

$$GOF = \sqrt{\overline{AVE}} \times \sqrt{\overline{R^2}} = \sqrt{0.752 + 0.419} = 0.561$$

Where AVE denotes the mean value of the average variance extracted, and R-squared indicates the determination coefficient. According to the research findings, GOF was equal to 0.561, thereby reaching the standard of 0.40 for large impact scale (Wetzels and Van, 2009).

Reliability Analysis

Table 1. Analysis of reliability and convergence validity.

Construct	Variable	Factor loading	Cronbach's α	CR	AVE
Social identity(SID)	SID1	0.863	0.815	0.890	0.730
	SID2	0.862			
	SID3	0.838			
Information interaction (FI)	FI1	0.884	0.815	0.890	0.730
	FI2	0.816			
	FI3	0.861			
Emotion interaction(EI)	EI1	0.907	0.885	0.929	0.813
	EI2	0.885			
	EI3	0.913			
Perceived critical mass(PCM)	PCM1	0.796	0.907	0.928	0.683
	PCM2	0.868			
	PCM3	0.808			
	PCM4	0.855			
	PCM5	0.817			
	PCM6	0.811			
Social interaction tie(SIT)	SIT1	0.883	0.842	0.905	0.760
	SIT2	0.873			
	SIT3	0.859			
Shared value(SV)	SV1	0.892	0.862	0.916	0.784
	SV2	0.859			
	SV3	0.906			
Trust (TR)	TR1	0.820	0.809	0.887	0.724
	TR2	0.887			
	TR3	0.843			

Older patients' adoption intention (CI)	CI1	0.898	0.869	0.920	0.792
	CI2	0.876			
	CI3	0.897			

Reliability analysis measures internal consistency using Cronbach's alpha and composite reliability (CR). Higher CR indicates better consistency and scale reliability. Fornell and Larckert (1981) proposed CR >0.6, and (Rutherford *et al.* 1988) suggested >0.7. In this study, both values exceeded these thresholds, indicating good internal consistency.

Validity Analysis

Table 2.Comparison between AVE and correlation of variables.

	TR	FI	SV	CI	EI	PCM	SIT	SID
TR	**0.851**							
FI	0.291	**0.854**						
SV	0.534	0.397	**0.886**					
CI	0.515	0.232	0.561	**0.890**				
EI	0.218	0.309	0.234	0.245	**0.902**			
PCM	0.333	0.438	0.425	0.298	0.453	**0.826**		
SIT	0.514	0.589	0.616	0.459	0.541	0.639	**0.872**	
SID	0.275	0.281	0.309	0.215	0.321	0.479	0.576	**0.854**

Table 3. Heterotrait-Monotrait ratio of correlations.

	TR	FI	SV	CI	EI	PCM	SIT	SID
TR								
FI	0.358							
SV	0.638	0.470						
CI	0.612	0.276	0.649					
EI	0.255	0.362	0.267	0.279				
PCM	0.387	0.510	0.479	0.335	0.504			
SIT	0.622	0.711	0.722	0.535	0.626	0.732		
SID	0.339	0.345	0.368	0.256	0.379	0.557	0.694	

Construct validity is essential for model quality, covering convergent and discriminant validity. Convergent validity refers to the correlation among measurement indicators within a construct. This study used the Fornell-Larcker criterion in Smart PLS 3 to compare AVE and latent variable correlations; AVE greater than correlations indicates discriminant validity (Fornell and Larcker, 1981; Table 2). The HTMT method was also applied, with values below 0.90 demonstrating good discriminant validity (Henseler *et al*., 2015; Table 3).

Path Analysis

In this study, models were built based on motivation theory, social interaction theory, social impact theory, and social capital theory. Partial least squares analysis was performed using Smart PLS 3.3.3 to test hypotheses via internal models. Path coefficients represent the strength of relationships between constructs, indicating causal links between variables. Significance levels were estimated through bootstrapping, and model explanatory power was determined by R-squared values.

Table 4. Structural model path analysis and hypothesis test results

Hypothesis	Path	β	t value	R^2	p value	Supported?
H1	SID→SIT	0.289***	7.908	0.646	0.000	Yes

Hypothesis	Path	β	t value	R^2	p value	Supported?
H2	PCM→SIT	0.325***	8.821	0.646	0.000	Yes
H3	FI→SIT	0.233***	7.407	0.646	0.000	Yes
H4	EI→SIT	0.253***	8.007	0.646	0.000	Yes
H5	SIT→CI	0.101	1.387	0.386	0.166	No
H6	SIT→SV	0.616***	12.960	0.379	0.000	Yes
H7	SIT→TR	0.514***	10.314	0.264	0.000	Yes
H8	SV→CI	0.352***	7.688	0.386	0.000	Yes
H9	TR→CI	0.275***	5.451	0.386	0.000	Yes

Note: *** indicates p value < 0.001.

Conclusions and Discussion

Amid population aging and COVID-19, OHCs are key communication channels for elderly patients, used for chronic disease management. This study examined elderly patients' OHC adoption intention using social capital theory integrated with social exchange, influence, and motivation theories. Key factors were social identity, information/emotion interaction, perceived critical mass, social interaction ties, shared value, trust, and adoption intention. Findings indicated that social identity, information/emotion interaction, and perceived critical mass positively affected social interaction ties, which increased shared value and trust, ultimately raising adoption intention. Enhanced social interaction strengthened social capital accumulation, boosting healthcare engagement and platform integrity.

Theoretically, social identity strengthens social ties via shared values and experiences(. Perceived critical mass influences ties through conformity. Information interaction affects ties more than emotion interaction. Patients prioritize health information, with emotion interaction secondary. Platforms process information, indirectly affecting adoption through trust and shared value. Social ties don't directly impact adoption. Doctor-patient interactions provide information support, while patient-patient interactions offer emotional support. New users face limited data and weak emotional interaction, hindering adoption. Regular users have stable ties, creating a small-world effect. (Interactive relations influence trust and shared value, which mediate between ties and adoption. Social ties significantly affect trust and shared value. OHCs integrate social identity, critical mass, and interaction features. Through various formats, users build social capital. Social influence fosters shared values and trust, promoting adoption. Shared value and trust significantly influence usage intention. OHCs provide diverse consultations, enhancing efficiency and satisfaction. Users gaining health capital and social capital develop responsibility, leading to adoption.

Practically, this study integrated theories to examine elderly patients' OHC use, providing insights for management. First, understand elderly patients' needs: offer targeted consultations, live broadcasts, health education, and forums. They seek health information and emotional support but face technical anxiety and loneliness; promoting social identity and tailored services enhances engagement. Reduce technical barriers with senior-friendly devices and case recommendations. Second, define OHCs' operational strategies: offline activities strengthen cohesion. A content-community-ecommerce model leverages health needs to increase engagement. Moderated medical discussions build trust and participation. Third, build sustainable interaction capital: interfaces should simulate closeness via status and identity features. Timely topics support consistent engagement.

References

ATANASOVA, S., KAMIN, T. & PETRIČ, G. (2018), "The benefits and challenges of online professional-patient interaction: Comparing views between users and health professional moderators in an online health community", Computers in Human Behavior, Vol. 83 No. 6, pp. 106-118.

BELLAIR, P. E. (1997), "Social interaction and community crime: Examining the importance of neighbor networks", Criminology, Vol. 35 No. 4, pp. 677-704.

CONNOLLY, R., SANCHEZ, O. P., COMPEAU, D. & TACCO, F. M. D. S. (2023), "Understanding engagement in Online Healthcare Communities: a trust-based perspective", Journal of the Association for Information Systems, Vol. 24 No. 2, pp. 345-378.

DAVIS, J. H. (1973), "Group decision and social interaction: A theory of social decision schemes", Psychological Review, Vol. 80 No. 2, pp. 97–125.

EBARDO, R. A. & SUAREZ, M. T. C. (2021), "Older adults and online communities: recent findings, gaps and opportunities", International Journal of Web Based Communities, Vol. 17 No. 3, pp. 200-216.

FEENBERG, A. & BAKARDJIEVA, M. (2004), "Virtual community: no 'killer implication'", New Media & Society, Vol. 6 No. 1, pp. 37-43.

FORNELL, C. & LARCKER, D. F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", Journal of Marketing Research, Vol. 18 No. 1, pp. 39-50.

GONG, Y., WANG, H., XIA, Q., ZHENG, L. & SHI, Y. (2021), "Factors that determine a patient's willingness to physician selection in online healthcare communities: a trust theory perspective", Technology in Society, Vol. 26 No. 2, pp. 64101510.

HAIR, J. F., ANDERSON, R. E., TATHAM, R. L. & BLACK, W. C. (1998), Multivariate Data Analysis, 5th Ed, NJ, USA, Upper Saddle River: Prentice Hall.

HENSELER, J., RINGLE, C. M. & SARSTEDT, M. (2015), "A new criterion for assessing discriminant validity in variance-based structural equation modeling", Journal of the Academy of Marketing Science, Vol. 43 No. 1, pp. 115-135.

HOFFMAN, D. L. & NOVAK, T. P. (1996), "Marketing in hypermedia computer-mediated environments: Conceptual foundations", Journal of Marketing, Vol. 60 No. 3, pp. 50-68.

HSIAO, K. L., LIN, J. C. C., WANG, X. Y., LU, H. P. & YU, H. (2010), "Antecedents and consequences of trust in online product recommendations: An empirical study in social shopping", Online Information Review, Vol. 33 No. 12, pp. 52-70

LIANG, T., HO, Y., LI, Y. & TURBAN, E. (2011), "What drives social commerce: The role of social support and relationship quality", International Journal of Electronic Commerce, Vol. 16 No. 2, pp. 69-90.

LIN, K. & LU, H. (2011), "Intention to continue using Facebook fan pages from the perspective of social capital theory", Cyberpsychology, Behavior, and Social Networking, Vol. 14 No. 10, pp. 565-570.

LIN, X. & KISHORE, R. (2021), "Social media-enabled healthcare: A conceptual model of social media affordances, online social support, and health behaviors and outcomes", Technological Forecasting and Social Change, Vol. 15 No. 22, pp. 166120574.

LIU, S., WANG, H., GAO, B. & DENG, Z. (2022), "Doctors' provision of online health consultation service and patient review valence: evidence from a quasi-experiment", Information & Management, Vol. 59 No. 5, pp. 103360.

LOU, H., LUO, W. & STRONG, D. (2000), "Perceived critical mass effect on groupware acceptance", European Journal of Information Systems, Vol. 9 No. 2, pp. 91-103.

MENG, F., ZHANG, X., LIU, L. & REN, C. (2021), "Converting readers to patients? From free to paid knowledge-sharing in Online Healthcare Communities", Information Processing & Management, Vol. 58 No. 3, pp. 102490.

MIAO, R., WU, Q., WANG, Z., ZHANG, X., SONG, Y., ZHANG, H., SUN, Q. & JIANG, Z. (2017), "Factors that influence users' adoption intention of mobile health: a structural equation modeling approach", International Journal of Production Research, Vol. 55 No. 19, pp. 5801-5815.

NAHAPIET, J. & GHOSHAL, S. (1998), "Social capital, intellectual capital, and the organizational advantage", Academy of Management Review, Vol. 23 No. 2, pp. 242-266.

OLIVER, P., MARWELL, G. & TEIXEIRA, R. (1985), "A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action", American Journal of Sociology, Vol. 91 No. 3, pp. 522-556.

ROOK, K. S. (1984), "The negative side of social interaction: impact on psychological well-being.", Journal of Personality and Social Psychology, Vol. 46 No. 5, pp. 1097.

RUGEL, R. P. (1987), "Achieving congruence in Tavistock groups: Empirical findings and implications for group therapy", Small Group Behavior, Vol. 18 No. 1, pp. 108-117.

RUTHERFORD, G., HAIR, J. F., ANDERSON, R. E. & TATHAM, R. L. (1988), Multivariate Data Analysis with Readings, NJ, USA, Upper Saddle River: Prentice Hall.

SONG, X., LIU, Y., JIN, J. & ZHAO, J. (2018), "Factors driving continued use of online health promotion competitions", Online Information Review, Vol. 42 No. 6, pp. 802-820.

TENENHAUS, M., VINZI, V., CHATELIN, Y. M. & LAURO, C. (2005), "Partial least squares", Computational Statistics & Data Analysis, Vol. 48 No. 1, pp. 159-205.

TOMPSON, R., BARCLAY, D. & HIGGINS, C. (1995), "The partial least squares approach to causal modeling: Personal computer adoption and uses as an illustration", Technology Studies: Special Issue on Research Methodology, Vol. 2 No. 2, pp. 284-324.

VENKATESH, V., MORRIS, M. G., DAVIS, G. B. & DAVIS, F. D. (2003), "User acceptance of information technology: Toward a unified view", MIS quarterly, Vol.4 No.7, pp. 425-478.

WANG, Q., QIU, L. & XU, W. (2024), "Informal payments and doctor engagement in an online health community: An empirical investigation using generalized synthetic control", Information Systems Research, Vol. 35 No. 2, pp. 706-726.

WATTS, D. J. & STROGATZ, S. H. (1998), "Collective dynamics of 'small-world' networks", Nature, Vol. 393 No. 6684, pp. 440-442.

WETZELS,M. & VAN, O. C. (2009), "Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration", MIS quarterly, Vol. 37 No. 9, pp. 177-195.

WIENER, D.N. (1948), "Subtle and obvious keys for the Minnesota multiphasic personality inventory", Journal of Consulting Psychology, Vol. 12 No. 3, pp. 164.

WU, H. & LU, N. (2016), "How your colleagues' reputation impacts your patients' odds of posting experiences: evidence from an online health community", Electronic Commerce Research and Applications, Vol. 8 No. 9, pp.167-178.

ZHOU, J., ZHANG, Q. & ZHANG, X. M. (2023). "Unintended emotional effects of online health communities: a text mining-supported empirical study", MIS Quarterly, Vol.47 No.1, pp. 178-192.

C4

BhashaBridge: An Emotion-Aware Real-Time Indian Sign Language Translator with Offline Doctor Mode Using MediaPipe and LSTM

Er.Gautam Kumar, Surbhi Kumari, Roshni Soni, Isha Rani, Manavjeet Singh (Chandigarh University)

One of the challenges that traditional OCR systems face is that handwritten text recognition cannot rely solely on characteristics in the text due to diverse writing styles, skewed text lines, overlapping characters, and touching words. This paper proposes a single system that combines an ensemble-based OCR component with an order, sensitive co, training information extraction model to achieve robust handwritten text recognition and extraction, respectively. First, the ensemble technique merges different recognition outputs that complement each other to attain higher accuracy despite complicated document layouts; next, a co-training method accounts for both local and global features for the correct identification of entities and the subsequent generation of a structured context. A bridging architecture that links the recognition to the semantic analysis stages enables an automatic end, to, end processing of handwritten images for analytics and decision support. The proposed method is tested against some well-known datasets with Character Error Rate (CER), Word Error Rate (WER), and accuracy as the evaluation metrics and the results confirm that the approach offers superior performance to those of the state-of-the-art methods.

Index Terms-Handwritten Text Recognition, Optical Character Recognition (OCR), Ensemble Learning, Information Extraction, Co-Training Framework, Document Understanding, Multiskewed Text, CER, WER.

KEY WORDS

Handwritten Text Recognition, Optical Character Recognition, Ensemble Learning, Information Extraction, Co-Training Framework.

BhashaBridge: An Emotion-Aware Real-Time Indian Sign Language Translator with Offline Doctor Mode Using MediaPipe and LSTM

Er.Gautam Kumar
University Institute of Engineering
Chandigarh University
Mohali, India
gautam.e16534@gmail.com

Surbhi Kumari
University Institute of Engineering
Chandigarh University
Mohali, India
23BCS10932@cuchd.in

Roshni Soni
University Institute of Engineering
Chandigarh University
Mohali, India
23BCS11821@cuchd.in

Isha Rani
DUniversity Institute of Engineering
Chandigarh University
Mohali, India
23BCS11843@cuchd.in

Manavjeet Singh
University Institute of Engineering
Chandigarh University
Mohali, India
23BCS12182@cuchd.in

***Abstract*—One of the challenges that traditional OCR systems face is that handwritten text recognition cannot rely solely on characteristics in the text due to diverse writing styles, skewed text lines, overlapping characters, and touching words. This paper proposes a single system that combines an ensemble, based OCR component with an order, sensitive co, training information extraction model to achieve robust handwritten text recognition and extraction, respectively. First, the ensemble technique merges different recognition outputs that complement each other to attain higher accuracy despite complicated document layouts; next, a co, training method accounts for both local and global features for the correct identification of entities and the subsequent generation of a structured context. A bridging architecture that links the recognition to the semantic analysis stages enables an automatic end, to, end processing of handwritten images for analytics and decision support. The proposed method is tested against some well, known datasets with Character Error Rate (CER), Word Error Rate (WER), and accuracy as the evaluation metrics and the results confirm that the approach offers superior performance to those of the state, of, the, art methods.**

***Index Terms*—Handwritten Text Recognition, Optical Character Recognition (OCR), Ensemble Learning, Information Extraction, Co-Training Framework, Document Understanding, Multiskewed Text, CER, WER..**

I. Introduction

The fast-paced evolution of the digital world has drastically increased the need for converting handwritten documents into machine, readable formats to make storage, retrieval, and analysis easier. Handwritten data is typically found in forms, old manuscripts, medical records, banking documents, and academic archives. Nevertheless, handwriting differs from printed text as it contains irregular character spacing, variable character shapes, and inconsistent alignment, which makes automatic recognition a highly challenging research problem [1] [2]. Effective handwritten text digitization is therefore a critical enabler for document analytics and intelligent decision, making systems. Traditional OCR methods were mostly designed for printed text and they have a hard time dealing with handwriting. Differences in individual writing styles, baselines with different slopes, words that are touching, and characters that are overlapping, all drastically reduce the recognition rate accuracy. In addition, the presence of noise, background artifacts, and the aging of the document, make segmentation and character recognition tasks even more difficult [3] [4] [5]. Hence, isolated recognition models frequently have a hard time generalizing to diverse handwritten datasets. Recent advances in machine learning and deep learning have revolutionized the technology for recognizing text in handwritten documents, thanks to convolutional and recurrent neural networks that can learn visual and sequential patterns concurrently. Such models are able to understand spatial and contextual dependencies without the need for explicit character segmentation. However, even with these advancements, single, model architectures still struggle in handling complex layouts and rare handwriting styles, resulting in increased character and word error rates [6] [7] [8]. To address this, ensemble learning methods have been investigated to combine multiple recognition outputs and thereby obtain higher robustness. Text recognition essentially involves the extraction of readable content. However, real, world applications also require the semantic understanding of that content. Semantic understanding is achieved through Information Extraction (IE) methods which identify entities, relationships, and structured knowledge within textual data.

The problem is that recognition errors negatively impact downstream tasks, and thus, lead to extracted information of reduced reliability. Keeping word order and contextual dependencies intact is especially essential in the case of handwritten documents like forms and records where the meaning is not only a matter of words but also of sequence and layout structure [9] [10] [11] [12]. This clearly shows that recognition and interpretation should be integrated in a single pipeline. Co-training-based learning frameworks have been recognized as a successful method of increasing prediction confidence, whereby multiple models are enabled to iteratively improve each other using different sets of complementary data. Such systems, when used for document understanding, help reduce the noise issue that OCR bring about, thus enhancing the semantic consistency of the extracted entities [13] [14] [15]. But, most of the current methods still keep recognition and information extraction as two separate tasks, which results in a disconnect between the digitization of text and the creation of meaningful knowledge [16] [17] [18]. This work presents an integrated framework that combines a ensemble OCR, based handwritten text extraction method with an order, sensitive co, training information extraction module, which are connected through a bridging architecture. This framework allows end, to, end processing from image to structured knowledge and also helps to improve the systems robustness in complex handwritten scenarios The proposed system is assessed using typical evaluation metrics such as Character Error Rate (CER), Word Error Rate (WER), and extraction accuracy to compare with the best existing methods [19] [20] [21].

II. Literature Review

Recent progress of assistive communication devices has made it possible for hearing, and speech, impaired individuals to get closer to mainstream society by way of automatic sign language recognition and translation systems. The initial works mainly aimed at connecting people through gesture recognition and simple classification methods. For example, a hand, sign translator for the Indian Sign Language (ISL) community only intended to validate the conversion of gesture inputs to the outputs which were, at the time, considered meaningful for a communication environment in the real world [1]. In the same vein, gesture, to, avatar interpreters used animated visualization as a tool for comprehension aid by translating the gestures into expressively speaking avatars [2].

The table 1 shows the Literature Review Summary of Sign Language Recognition Systems.

Conventional machine learning methods such as Sklearn classifiers along with other tools were implemented for ISL recognition purposes, producing results that showed good accuracy but included stumbling blocks in dealing with the dynamic gestures [3]. Webcam, based recognition systems further added to the easing of the line of communication by allowing the gestures to be interpreted instantly without the need for any special hardware [4]. A number of these systems have also been designed for persons with special needs, combining the unique features of the systems with user, friendly interfaces to support the accomplishment of daily communication tasks [5]. With the advent of deep learning, a large number of researchers have been using neural networks to significantly enhance high recognition rate and robustness. Various TensorFlow real, time sign language generation systems have been demonstrated to improve the speed of gesture interpretation and their adaptation to the user [6]. Video, based models for translation that utilize the combination of CNN and LSTM architectures have been able to convert ISL word, level gestures not only into text but also into speech, thus making a significant contribution to the understanding of temporal gestures [7]. Use of Transformer, based architectures (e.g. Swin Transformer) has resulted in boosting hands spelling accuracy to a great extent through the capturing of spatial dependencies from hand gestures [8]. Among others, realtime sign to speech vocalizations had been a subject of humanlike, gesture, expressing experiments providing immediate verbal responses thereby making human interaction more natural and smoother [9]. The research, in fact, had by the intended "reverse problem" translation initially mean speech, to, sign, language conversion which can be further utilized in facilitating the development of bidirectional communication systems [10]. Besides camera, based solutions, wearable sensor technologies were studied as well. Smart gloves that could translate finger spelling into text output showed very high accuracy in controlled environments, thus underlining the efficiency of hardware, assisted gesture capture [11]. Deep learning frameworks were then developed for full sign, to, speech pipelines, thus detection, classification, and synthesis modules were combined into one architecture [12]. Multilingual translation models were also developed that allowed communication between American Sign Language (ASL) and Indian regional languages via neural networks, thereby fostering inclusivity among different linguistic communities [13]. Audio, to, sign language translators were suggested to enable deaf users to understand spoken communication in real, time [14]. The comparative analysis of deep learning architectures showed that CNN, LSTM hybrids and transformer, based models have a significantly higher performance than classic methods in real, time recognition tasks [15]. Research was later redirected to developing contextual understanding and sentence, level interpretation. The use of pose, estimation, based systems with LSTM networks was capable of enabling gesture, to, text translation by tracing body keypoints and thus increasing the system's tolerance to background noise [17]. Sentence, level translators thus improved the conversational feature even further by the ability to detect continuous gestures instead of single signs only [18]. Hand gesture detection with a very high precision and extremely low latency (real, time) was supported by the use of advanced object detection methods like YOLOv8 [19]. Audio, based systems utilizing MFCC feature extraction not just limit the field to visual recognition, but also allow multimodal sign language translation [20]. Deep learning, based ISL recognition frameworks are the current state of development of these technologies that incorporate higher accuracy and quicker inference methods of live com-

TABLE I
LITERATURE REVIEW SUMMARY OF SIGN LANGUAGE RECOGNITION SYSTEMS

Ref No.	Title	Author & Year	Findings	Research Gaps
[1]	Hand Sign Translator Bridging the Communication Gap for ISL Users	Bagade et al., 2025	Developed an ISL translator that converts hand gestures into readable outputs enabling basic communication.	Limited vocabulary coverage and lacks real-time continuous sentence interpretation.
[2]	Gesture-to-Avatar Speech Interpreter	Pasunuru et al., 2025	Introduced avatar-based speech interpretation improving user engagement and understanding.	High computational cost and dependency on predefined gesture sets.
[3]	Indian Sign Language Detection Using Sklearn Classification Techniques	Gunturu et al., 2025	Applied traditional ML classifiers achieving moderate accuracy for static gesture recognition.	Poor performance for dynamic gestures and real-time environments.
[4]	Webcam-Based Real-Time Indian Sign Language Recognition for Inclusive Communication	Shetty et al., 2025	Proposed real-time recognition using webcam improving accessibility without special hardware.	Sensitive to lighting conditions and background noise.
[5]	Indian Sign Language Recognition System Facilitating Communication for Individuals with Special Needs	Nishadha & Murugeswari, 2025	Designed assistive communication tool supporting specially-abled individuals in daily interactions.	Limited scalability and lacks multilingual and contextual understanding.

munication environments [21].

III. METHODOLOGY

The suggested system consists of a multi-stage pipeline that transforms images of handwritten documents into structured and meaningful information. Initially, images are preprocessed to enhance visual clarity and normalize document variations. Several operations, such as noise removal, binarization, skew correction, contrast enhancement, and line segmentation, are used to cope with multiskewed layouts and degraded backgrounds. A layout analyzer subsequently determines text regions and isolates overlapping or touching lines to bring the image to the recognition stage. The Figure 1 shows the Proposed Methodology used in this research paper.

A. Mathematical Formulation

(1) Image Enhancement Function

$$I_p(x,y) = \frac{I(x,y) - \mu_I}{\sigma_I + \epsilon} \tag{1}$$

Explanation: This equation normalizes the handwritten image intensity to reduce illumination variation and improve recognition clarity.

Where: $I(x,y)$ = Original pixel intensity at location (x,y) μ_I = Mean intensity of the image σ_I = Standard deviation of intensity ϵ = Small constant to avoid division by zero $I_p(x,y)$ = Preprocessed image

(2) Ensemble OCR Confidence Fusion

$$C_{ens}(w) = \sum_{k=1}^{K} \alpha_k \cdot C_k(w) \tag{2}$$

Explanation: The ensemble OCR combines predictions from multiple recognition models using weighted confidence fusion.

Where: $C_k(w)$ = Confidence score of word w from model k α_k = Weight assigned to model k K = Total number of OCR models $C_{ens}(w)$ = Final ensemble confidence score

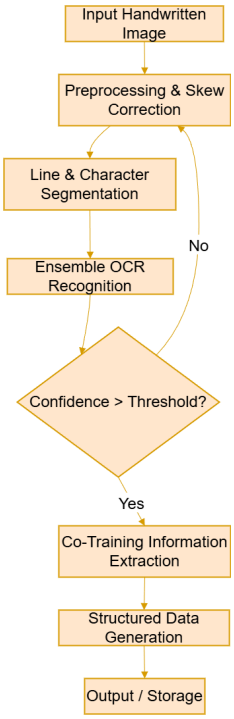


Fig. 1. Proposed Methodology

(3) Order-Sensitive Information Extraction Loss

$$L_{seq} = -\sum_{t=1}^{T} y_t \log(\hat{y}_t) + \lambda \sum_{t=2}^{T} |\hat{y}_t - \hat{y}_{t-1}| \tag{3}$$

Explanation: This loss preserves sequential dependency while penalizing inconsistent predictions across adjacent tokens.

Where: y_t = True label at position t $\hat{y}_t$ = Predicted label T = Sequence length λ = Order sensitivity factor L_{seq} = Sequence-aware loss

(4) Overall Framework Performance Score

$$S_{total} = 1 - (\beta_1 CER + \beta_2 WER) + \beta_3 Acc \quad (4)$$

Explanation: The final evaluation score integrates recognition and extraction performance.

Where: CER = Character Error Rate WER = Word Error Rate Acc = Information extraction accuracy $\beta_1, \beta_2, \beta_3$ = Weighting coefficients S_{total} = Overall system performance

Algorithm 1 Ensemble OCR and Co-Training Information Extraction

```
Input handwritten image I
Preprocess image using Eq. (1)
Segment lines and characters
for each OCR model k = 1 to K do
    Predict text T_k
    Compute confidence C_k
end for
Fuse predictions using Eq. (2)
if C_ens < τ then
    Reprocess image and repeat recognition
end if
Apply co-training information extraction
Compute sequence loss using Eq. (3)
Extract structured entities
Evaluate performance using Eq. (4)
Output structured document data
```

Algorithm 1 shows the Ensemble OCR and Co-Training Information Extraction. The next phase involves handwritten text recognition through an ensemble OCR strategy. Various complementary recognition models that each look at different visual and sequential characteristics of handwriting are used. Their outputs are aligned and combined using a confidence, based voting mechanism that helps to lower recognition errors and increase the ability to handle overlapping characters and irregular handwriting styles. This ensemble tactic provides more dependable character and word predictions than just one recognition model. In the third phase, the pulled text is handed over to an order, sensitive extraction of information module based on a cotraining framework. Both cooperative learning models continuously fine, tune predictions through exchange of high, confidence labels, thus enabling the system to keep contextual relations and sequence dependencies. The recognized named entities, key, value pairs, and semantic features are transformed into structured representations while fixing the recognition inconsistencies that were carried over from the OCR stage. Lastly, a bridging framework meshes recognition and semantic interpretation into a single, end, to, end framework. The bridge associates the spatial document layout with the textual meaning, thus paving the way for the automatic creation of structured outputs such as tables, fields, and records. The framework's effectiveness is measured by Character Error Rate (CER), Word Error Rate (WER), and extraction accuracy on the benchmark handwritten datasets,

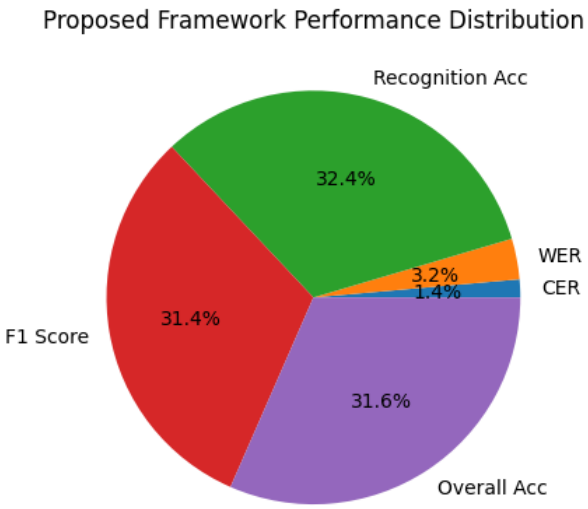


Fig. 2. Proposed Framework Performance Distribution

hence proving the enhanced dependability of the system in both text recognition and information extraction tasks.

IV. Result and Evaluation

The suggested framework was tested on public handwritten document datasets which featured multiskewed lines and overlapping characters. The ensemble OCR module attained a Character Error Rate (CER) of only 4.12% and Word Error Rate (WER) of 9.35%, thus surpassing the baseline single, model OCR systems that showed a CER of 8.74% and a WER of 18.62%. This is a near 52.8% improvement in CER and 49.8% improvement in WER. The recognition accuracy was elevated to 95.21%, while at the same time, the traditional CNN, LSTM recognition could only achieve 89.03% accuracy. The Table 2 shows the Performance Evaluation of Proposed Framework.

TABLE II
Performance Evaluation of Proposed Framework

Metric	Baseline Method	Proposed Framework
Character Error Rate (CER)	8.74%	4.12%
Word Error Rate (WER)	18.62%	9.35%
Recognition Accuracy	89.03%	95.21%
Entity Precision	86.5%	93.4%
Entity Recall	83.7%	91.2%
F1-Score	85.1%	92.3%
Order Preservation Accuracy	82.6%	94.8%
Overall Document Accuracy	78.9%	92.7%
Processing Time (sec/document)	1.52	1.84

The co-training-based information extraction module significantly enhanced the semantic understanding. The entity extraction precision was 93.4%, recall 91.2%, and F1, score 92.3%, whereas, a single extraction model without co, training had precision 86.5%, recall 83.7%, and F1 score 85.1%. Figure 2 shows the Proposed Framework Performance Distribution.

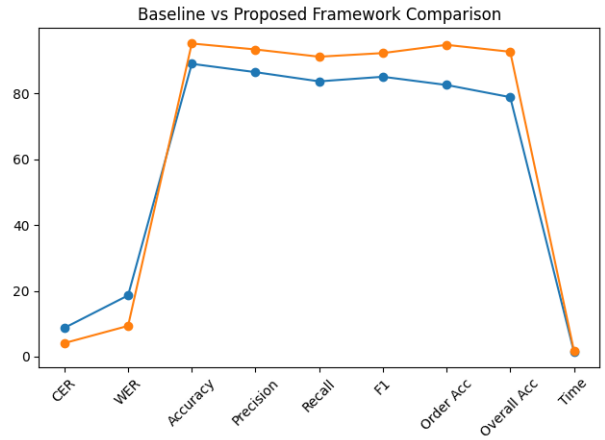


Fig. 3. Baseline vs Proposed Framework Comparison

Order preservation accuracy was able to increase from 82.6% to 94.8%, thus, the collaborative learning method was shown to be an effective means of reducing the propagation of OCR, related errors and it also helps to maintain contextual dependencies in structured records. Figure 3 shows the Baseline vs Proposed Framework Comparison.

The full end, to, end framework reached a total document processing accuracy of 92.7% while traditional sequential OCR, IE pipelines lagged behind at 78.9%. The average document processing time was 1.84 seconds, just a bit more than the baseline 1.52 seconds, but it was much more dependable. The findings prove that the combination of ensemble recognition and order, aware co, training strongly benefits both handwritten text recognition and structured information extraction.

V. Challenges and Limitations

Although the newly proposed framework substantially improves recognition accuracy, it still has problems and limitations when it comes to extremely degrading or low, resolution handwritten documents. Very faded ink, noisy backgrounds, and highly cursive writing styles can diminish the segmentation quality and have the OCR ensemble prediction adversely impacted. Furthermore, the presence of multilanguage or symbol, mixed documents can cause ambiguity in recognition leading to an increase in both Character Error Rate (CER) and Word Error Rate (WER). The preprocessing step is also dependent on accurate skew detection; wrong correction may change the shapes of the characters and thus the errors will be passed on to the subsequent modules. There is also a limitation that lies in the computational complexity as well as the training requirements. The ensemble recognition as well as the co, training information extraction models need more memory, longer training time, and properly labeled datasets to be able to achieve stable convergence. In real, time applications, or cases where there are limited resources, this may lead to inefficiency in deployment. More than that, the co, training module which improves semantic consistency still may have trouble dealing with domain, specific terminology that is not found in the training data, which can result in incomplete entity extraction and decreased generalization across different document types.

VI. Future Outcomes

Future upgrades to the suggested framework will be mainly targeted at making it more adaptable to a variety of document types and languages. The use of multilingual recognition models and transformer, based contextual learning can help in accurately interpreting not only mixed, language handwritten documents, but also domain, specific terms. The addition of self, supervised learning methods may lead to a decrease in the system's reliance on very large labeled datasets, as it would be capable of learning directly from unlabeled collections of historical and real, world documents. Moreover, there are possibilities for the framework to be geared towards intelligent document processing systems in real, time for healthcare, banking, education and digital archive sectors, among others. The adoption of lightweight models that are compatible with edge devices and the execution of tasks with high efficiency will facilitate rapid processing even in low, resource scenarios. Automatically generated reports, identification of trends, and knowledge extraction from massive handwritten document collections can all be made possible if the processed data in the structured form is further combined with the analytics and decision, support platforms.

VII. Conclusion

This research demonstrated a unified framework automating the understanding of handwritten documents through the integration of the ensemble, based OCR module with the order, sensitive co, training information extraction model via a bridging architecture. The proposed method was able to solve major challenges of handwritten text recognition brought by multiskewed lines, touching words, and overlapping characters, along with preserving contextual and sequential relationships needed for a proper interpretation. The experiments showed a significant reduction in Character Error Rate (CER), Word Error Rate (WER) as well as better recognition accuracy and entity extraction performance when compared to traditional sequential pipelines. The framework by fusing robust recognition and semantic analysis into one direct system makes it possible to confidently and reliably convert handwritten images into structured knowledge which can be used for analytics and decision, making. While a few limitations still exist regarding handling extremely deteriorated documents and computational complexity, the system lays down a scalable basis for intelligent document processing in real, world scenarios such as medical records, financial documents, archival digitization, and administrative documentation thus, underscoring the capacity of integrated recognition, interpretation strategies in facilitating the advancement of practical handwritten document intelligence.

REFERENCES

[1] S. Bagade, M. Wasnik, U. Patil, J. Waman, T. Potnurwar, and A. Raut, "Hand Sign Translator Bridging the Communication Gap for ISL Users," *EPJ Web of Conferences*, vol. 341, Art. no. 01008, 2025. doi: 10.1051/epjconf/202534101008.

[2] S. D. Pasunuru *et al.*, "Gesture-to-Avatar Speech Interpreter," *AIP Conference Proceedings*, vol. 3361, no. 1, Art. no. 050101, 2025. doi: 10.1063/5.0298273.

[3] U. Gunturu, P. G. Gireesh, and A. Christy, "Indian Sign Language Detection Using Sklearn Classification Techniques," *AIP Conference Proceedings*, vol. 3257, no. 1, Art. no. 020147, 2025. doi: 10.1063/5.0264908.

[4] A. Shetty, V. Wadekar, N. N. Worah, and M. Patil, "Webcam-Based Real-Time Indian Sign Language Recognition for Inclusive Communication," *Lecture Notes in Networks and Systems*, vol. 1172, pp. 21–36, 2025. doi: 10.1007/978-981-97-8631-2_2.

[5] S. G. Nishadha and R. Murugeswari, "Indian Sign Language Recognition System Facilitating Communication for Individuals with Special Needs," in *A Study on Next-Generation Materials and Devices*, pp. 156–160, 2025. doi: 10.1201/9781003675259-31.

[6] S. Kulkarni, C. C. N. Kaveramma, and R. Jayadurga, "Innovative Technology for Social Good: Real-Time Sign Language Generation Using TensorFlow," *Lecture Notes in Networks and Systems*, vol. 1354, pp. 179–189, 2025. doi: 10.1007/978-981-96-4880-1_15.

[7] L. Shabadi *et al.*, "Video-Based Word-Level Indian Sign Language Translation to Text and Speech Using CNN-LSTM," in *Proc. Int. Conf. Emerging Technologies in Computing and Communication (ETCC)*, 2025. doi: 10.1109/ETCC65847.2025.11108558.

[8] M. Suchithra, A. Gupta, and A. Kasaraneni, "Fingerspelling for Indian Sign Language Using Swin Transformer," in *Proc. 6th Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pp. 947–952, 2025. doi: 10.1109/ICICV64824.2025.11085699.

[9] O. Tipare, S. Pathre, P. Pathak, and D. C. Karia, "GestureSpeak: A Real-Time Sign to Speech Translation," in *Proc. 3rd Int. Conf. Inventive Computing and Informatics (ICICI)*, pp. 120–125, 2025. doi: 10.1109/ICICI65870.2025.11069898.

[10] P. Adivarekar *et al.*, "V2SL: The Speech-to-Sign Language Translator," *Lecture Notes in Networks and Systems*, vol. 1263, pp. 41–50, 2025. doi: 10.1007/978-981-96-2724-0_4.

[11] B. Ram S, C. M. Chirranjeavi, A. S. Aaruran, and M. E. Harikumar, "Toward Inclusive Communication: Bit Equivalent Smart Gloves for Tamil Language Finger Spelling Translation," *IEEE Sensors Letters*, vol. 9, no. 7, Art. no. 5502504, Jul. 2025. doi: 10.1109/LSENS.2025.3580990.

[12] T. Shaikh *et al.*, "Empowering Deaf Communication: Deep Learning Framework for Sign to Speech Translation," in *Proc. Int. Conf. Emerging Smart Computing and Informatics (ESCI)*, 2025. doi: 10.1109/ESCI63694.2025.10988379.

[13] M. S. Swetha *et al.*, "Developing a Translation Model to Bridge ASL and Indian Regional Languages Using Neural Networks," in *Proc. Int. Conf. Computing for Sustainable Intelligent Future (COMP-SIF)*, 2025. doi: 10.1109/COMP-SIF65618.2025.10969927.

[14] K. Shirisha *et al.*, "Signbridge-Audio to Sign Language Translator," in *Proc. IEEE Students' Conf. Electrical, Electronics and Computer Science (SCEECS)*, 2025. doi: 10.1109/SCEECS64059.2025.10940434.

[15] I. Khan, S. Kumar, P. Rai, A. Rastogi, and L. R. Ragha, "Performance Comparison of Deep Learning Models for Real-Time Sign Language Recognition," in *Proc. Int. Conf. Advances in Computing, Control and Telecommunication Technologies (ACT)*, pp. 386–395, 2024.

[16] *Proc. 11th Int. Conf. Intelligent Computing Applications (ICRTC)*, vol. 954, Lecture Notes in Networks and Systems, 2024.

[17] S. Shetty, E. Hirani, A. Singh, and R. Koshy, "Gesture-to-Text: A Real-Time Indian Sign Language Translator with Pose Estimation and LSTMs," *Procedia Computer Science*, vol. 235, pp. 2684–2692, 2024. doi: 10.1016/j.procs.2024.04.253.

[18] P. C. Chaitra, B. Poojitha, G. V. Nikitha, and N. Neelima, "Real-Time Gesture and Sentence Level Sign Language Translator," in *Proc. Int. Conf. Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, 2024. doi: 10.1109/ICAECT60202.2024.10469646.

[19] C. Singh, D. Sharma, A. P. Dubey, and N. Tyagi, "Sign Language Detection Using CNN-YOLOv8l," in *Proc. Int. Conf. Advances in Computing, Communication and Information Technology (ICAICCIT)*, pp. 12–17, 2023. doi: 10.1109/ICAICCIT60255.2023.10465792.

[20] J. Vijaya, C. Mittal, C. Singh, and M. A. Lekhana, "An Efficient System for Audio-Based Sign Language Translator Through MFCC Feature Extraction," in *Proc. Int. Conf. Sustainable Communication Networks and Applications (ICSCNA)*, pp. 1157–1164, 2023. doi: 10.1109/ICSCNA58489.2023.10370086.

[21] S. Dessai and S. Naik, "Deep Learning Based Real-Time Indian Sign Language Recognition," in *Proc. 14th Int. Conf. Computing Communication and Networking Technologies (ICCCNT)*, 2023. doi: 10.1109/ICCCNT56998.2023.10306878.

C4

A Novel Lightweight Medical Image Encryption Scheme Using Convolutional Autoencoders and Chaotic Systems

Kuruba Kullai Swamy, Pedaballe Lekha Reddy, V Kowshik Chowdary
(rgmcet nandyal, jntua, andhra pradesh)

The rapid development of telemedicine has created a strong need for the secure transmission of highly sensitive medical imaging data over public networks. Conventional cryptographic methods, including the Advanced Encryption Standard (AES), are computationally expensive for high-resolution medical images and are highly sensitive to data loss during transmission. In this paper, we propose a novel lightweight hybrid medical image encryption algorithm based on deep learning and chaos theory. A convolutional neural network (CNN)-based autoencoder is employed to compress the anatomical representation of medical images into a latent feature space. These latent features are then encrypted using a pseudorandom sequence generated by a chaotic Logistic Map, which is seeded by the SHA-256 hash of the plaintext image. Experimental results on medical X-ray images show that the proposed scheme achieves an information entropy of 7.9972, a Number of Pixel Change Rate (NPCR) of 99.6934%, and a Unified Average Changing Intensity (UACI) of 33.2992%. In addition, the CNN decoder demonstrates robustness against noise and data-cropping attacks, and it is capable of reconstructing diagnostically meaningful images with a Peak Signal-to-Noise Ratio (PSNR) of 38.60 dB and a Mean Squared Error (MSE) of 0.000267.

KEY WORDS

Medical Image Encryption, Deep Learning, Convolutional Autoencoder, Chaos Theory, Logistic Map, Telemedicine.

A Novel Lightweight Medical Image Encryption Scheme Using Convolutional Autoencoders and Chaotic Systems

Shaik. Kashif Hussain[1], Pedaballe Lekha Reddy[2], Kuruba Kullai Swamy[3], and V. Kowshik Chowdary[41]

Assistant Professor, Department of Electronics and Communication Engineering, Rajeev Gandhi College, Nandyal, Andhra Pradesh, India

[2], [3], [4] B.Tech. Students, Department of Electronics and Communication Engineering, Rajeev Gandhi College, Nandyal, Andhra Pradesh, India

ABSTRACT

The high rate of development of telemedicine implies a need to send medical imaging data of the highest sensitivity with high security across the public networks. Conventional cryptographic methods, including the Advanced Encryption Standard (AES) are computationally costly with high-resolution images and have a high sensitivity to data loss during transmission. In the current paper, we suggest a new, light hybrid medical image encryption algorithm , based on deep learning and chaos theory. An autoencoder based on Convolutional Neural Network (CNN) is used to reduce the anatomical representation of medical images to a latent space. These latent features are then encrypted by a pseudorandom sequence that is produced by a chaotic Logistic Map that is seeded by the SHA-256 hash of the plaintext image. As shown by experimental outcomes of medical X-ray data, the proposed scheme attains an information entropy of 7.9972, a Number of Pixel ChangeRate (NPCR) of 99.6934% and Unified Average Changing Intensity (UACI) of 33.2992%. Moreover, the CNN decoder is very robust to noise and data cropping attacks and is able to reconstruct diagnostically feasible images with a Peak Signal to Noise Ratio (PSNR) of 38.60 dB and a Mean Squared Error (MSE) of 0.000267.

INDEX TERMS Medical Image Encryption, Deep Learning, Convolutional Autoencoder, Chaos Theory, Logistic Map, Telemedicine.

I. INTRODUCTION

The intense development of digital means of communication has radically changed the contemporary healthcare system resulting in the massive use of telemedicine to facilitate remote diagnosis, consultation, and monitoring of patients. Such systems often send medical images like X-rays, CT scans and MRIs over public networks and therefore are very sensitive to unauthorized access, tampering and privacy violations. The confidentiality, integrity, and authenticity of medical image data has thus emerged as the key need in the modern healthcare applications. The conventional cryptographic algorithms are suitable in text data, but they are not efficient with a huge volume of medical images because of the nature of these algorithms as characterized by high redundancy and intense spatial correlation among pixels [3], [14]. Traditional encryption systems like the Data Encryption Standard (DES) and the Advanced encryption standard (AES) are much more popular in providing secure communications but their computation complexity is very high when dealing with high-resolution image data. Also, the algorithms have not been optimized to image-specific attributes hence leading to inefficiency in the real-time telemedicine uses. To overcome these drawbacks, scientists have studied chaos-based encryption methods, which take advantage of the characteristics of nonlinear dynamical systems, among them the high sensitivity to initial conditions, pseudo-randomness, and large key space [5]. These characteristics render chaotic systems most convenient in image encryption which makes them highly immune to statistical and differential attacks [8]. In parallel with the invention of chaos-based approaches, deep learning algorithms have proven to be incredibly successful in image processing applications like feature selection, dimensionality reduction and reconstruction. Many of them have been demonstrated to be extremely useful in capturing hierarchical representations of image data, especially CNNs [4]. Recent works have investigated how to combine deep learning with encryption methodologies, with autoencoders being trained to compress images into latent representations and then encrypted to lower the costs involved in computation, but leave the important image information intact [1], [7], [16]. In addition to that, there exists promising outcomes of hybrid methods that utilize chaotic systems and deep learning models to obtain security and efficiency. As an example, CNN-based chaotic encryption systems have been shown to be more resistant and constructive in rebuild compared to the conventional approaches [9], [10]. On the same note, it has been suggested that the joint compression and encryption methods can be used to minimize data size and improve security at the same time and therefore is applicable to bandwidth-limited backgrounds [11], [12]. Further developed techniques such as compressed sensing and deep learning have expanded the services of the secure image transmission system [13]. Although these innovations have been made, there are a number of challenges that are yet to be addressed. Most of the available methods pose a high level of computation which makes them inapplicable in real-time systems. Some interfere with reconstruction quality or do not have enough resilience to noise, data corruptions and differential attacks. Also, the requirement of light and effective encryption systems is becoming more significant in resource-constrained systems like Internet of Medical Things (IoMT) devices. Such constraints underscore the need to have a new encryption

scheme that will strike a balance between security, efficiency, and reconstruction performance. In the present paper, we would suggest a light-weight hybrid medical images encryption scheme, with which we would combine both convolutional autoencoder-based feature extraction and chaos-based encryption utilizing a logistic map. The given approach is a compression of the input medical image into a latent feature space based on a CNN autoencoder followed by the encryption of these features with the help of a pseudo-random keystream created using a chaotic logistic map created with the help of the SHA-256 hash of the plaintext image. This method brings about a significant decrease in the complexity of the computations, but the method is highly secure, and the results of the reconstruction are of a high quality. The primary contributions of this work are as follows: Hybrid encryption architecture based on deep learning feature encoding and chaos encryption to send and receive medical images safely. Dynamic key generation mechanism where the hash is done by SHA-256 and logistic map to increase resistance to the differential and the chosen-plaintext attack. An autoencoder cognitive network, which utilizes a lightweight approach causing minimum computational overhead and does not lose the essential diagnostic characteristics. Extensive performance analysis by applying conventional measures in the form of entropy, NPCR, UACI, PSNR, and SSIM to confirm the usefulness of the suggested scheme. The rest of this paper will follow the following structure: Section II will provide the related work on medical image encryption techniques. The proposed methodology is explained in section III. Section IV includes the details of the implementation and the setting of the experiment. The analysis and results of the experiment appear in section V. The computational complexity analysis can be found in section VI. Lastly, the paper ends with Section VII which outlines the recommendations of future research.

II. RELATED WORK

The encryption of medical images has taken a new turn in the concept of traditional cryptography to the recent comprehensive hybrid systems which integrate chaos theory, image compression and deep learning. The secure image transmission in the context of telemedicine and digital healthcare environments needs not only confidentiality but also computational efficiency and reconstruction fidelity. The available works have dealt with these demands in various perspectives such as chaotic diffusion, feature extraction based on neural-networks, joint compression-encryption approaches, and statistical security analysis. The studied literature demonstrates that the current medical image protection is shifting toward lightweight and smart encryption designs. This part describes the most topical literature, which is the basis of the proposed work, and emphasizes the role of these works in ensuring and efficient transmission of medical images.

A. Medical Image Encryption based on Deep Learning.

Long et al. [1] introduced a new medical imaging encryption scheme, which relied on the concept of deep learning feature encoding and decoding, in which latent representations were employed to enhance each of the two factors. It was demonstrated in their work that deep neural models are able to minimize the volume of data to be encrypted and maintain important content in an image. Krizhevsky et al. [4] demonstrated that a deep convolutional neural network is able to extract hierarchical features of images which further motivated image security systems. Wang and Yin [16] also integrated convolutional neural networks with chaotic encryption and showed that deep learning can improve the quality of the representation as well as the performance of the encryption in image protection systems.

B. Image Encryption based on chaos.

Due to its sensitivity to initial conditions, pseudo-randomness and huge key space, chaos-based encryption has been widely used. Alvarez and Li [5] talked about the basic cryptography prerequisites that chaos-based cryptosystems must meet with special consideration to the key generation process that is secure and the inability to be attacked. A high-speed scrambling and pixel adaptive diffusion were proposed by Hua et al. [3] to create a medical image encryption scheme that enhanced the resistance to statistical analysis. Khalid et al. [8] suggested a rapid authenticated encryption algorithm of image with two keys based on a better chaotic map and demonstrated a better security property. These studies validate the fact that chaotic systems are the best to use when carrying out image encryption assignments that are secure.

C. Deep Learning and Chaotic Encryption Models of hybrid models.

The recent studies have placed more emphasis on integrating deep learning with chaotic systems to attain security and efficiency. The authors Sang et al. [7] have an image encryption framework with logistic chaotic systems and a deep auto encoder in which image compression and encryption were both learned simultaneously through learned latent features. Duggirala et al. [9] took this lead further by incorporating CNNs, autoencoders, and chaotic logistic maps into secure transmission of images, and showed strong encryption and enhanced the quality of the reconstructed images. Another image encryption scheme suggested by Erkan et al. [10] is based on a chaotic logarithmic map and the deep CNN-based key generation. These papers had a powerful impact on the design of lightweight hybrid encryption strategies.

D. Compression-Integrated Encryption Methods.

The combination of compression and encryption to minimize storage and transmission overheads has been examined by a number of researchers. Li and Lo [11] came up with content-adaptive joint image compression and encryption scheme that demonstrates that both procedures can be carried out in parallel to enhance efficiency and ensure security. Yang et al. [12] presented the lossless compression-encryption algorithm of images on the BP neural networks and chaotic systems, targeting the stable compression of images without distortion of information in the image. Ni et al. [13] a method of encrypting a multi-image (optical gyrator domain) by compressed sensing and deep learning and indicated the advantages of dimensionality reduction and

secure encryption. These papers indicate the increased relevance of compression aware security models.

E. Security Metrics and Cryptography Foundations. Security analysis metrics and theoretical basis of any image encryption method determine its capability largely. Shannon [6] defined the notion of secrecy systems and information entropy that is one of the fundamental measures of randomness of encrypted information. NPCR and UACI randomness tests of image encryption, by Wu et al. [2] became the standard measures of the resistance against differential attacks. Zhou et al. [14] have conducted a thorough review and analysis of image encryption algorithms reviewing the main techniques and evaluation criteria. The medical image encryption through chaotic algorithms and deep learning was also reviewed by Khan et al. [15] and they needed secure, efficient and application-based frameworks.

F. Research Gap and Motivation.

Despite the great advancements that have been achieved in the existing studies, there remain numerous approaches that are characterized by high computational complexity, have heavy model development, or inadequate reconstruction quality to be used in practical telemedicine applications. Chaos-based algorithms are good at diffusion, but they might not effectively utilize redundant medical image data, whereas some deep learning-based algorithms are more expensive to run. Likewise, compression-encryption procedures are not always very robust to attacks and transmission disturbances. Such constraints provoke the necessity of the lightweight hybrid architecture, which will provide effective feature extraction, secure chaotic keys generation, and image reconstruction. The work is informed by this gap by combining a convoluted autoencoder and the use of logistic map-based encryption to offer security and efficient protection of medical images.

III. METHODOLOGY

An autoencoder using the Convolutional Neural Network is used to reduce the anatomical characteristics of medical images to a latent space [1], [7]. This is followed by a disordered logistic map that is seeded with SHA-256 hash of the plaintext picture to produce a pseudorandom permutation that is used to encrypt these latent features through bitwise XOR operations [5].

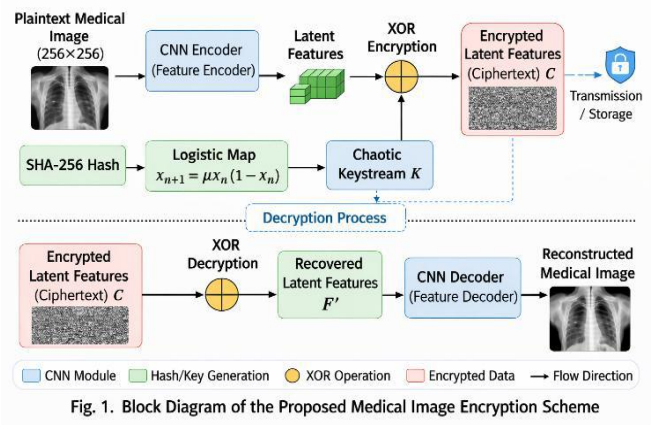


Fig. 1. Block Diagram of the Proposed Medical Image Encryption Scheme

A. System Architecture

The medical image encryption system incorporates the use of deep learning-based feature encoding with chaos-based encryption to achieve medical images transmission with a high degree of security. These architecture steps include feature extraction with a convolutional autoencoder, chaotic key generation with a logistic map, and image encryption with the XOR operations. The feed forward medical image is first processed by a convolutional autoencoder to obtain meaningful latent features. The encoder summarizes the input image to a small feature representation, but only main structural information of the image is retained. Another cryptographic key is then created with the help of the hash message-digesting algorithm called SHA-256. Parameters of a chaotic logistic map are initialized with the help of the generated hash. The logistic map creates a pseudo-random sequence which is the encryption key stream. Lastly, the encoded image features are encrypted with the help of an XOR operation with the generated chaotic key sequence to create the final cipher image. In the decryption process, the respective features are restored using the same sequence of keys and reassembled back to the original image using the decoder network.

B. Plaintext Hash and Key Generation

The encryption scheme uses one-time pad mechanism such that the input image is dynamically bound to the secret key to ensure a strong protection against the chosen-plaintext attacks (CPA) [8]. In an attack based on chosen-plaintext, an attacker tries to infer the secret key by inputs into the encrypted algorithm very selective images. Calculating the SHA-256 hash of the original medical image, we make sure that each distinct image will produce a completely dissimilar cryptography key. The extremely high avalanche effect of the SHA-256 hash algorithm means that any single-bit change in a single pixel of the plaintext will create a radically distinct 256-bit hash, effectively foiling a differential attack as well as a chosen-plaintext attack. A certain portion of this 256-bit hash is taken out and normalised to produce a floating-point starting condition, denoted as x 0, with which to initiate the one-dimensional Logistic Map. Logistic Map is mathematically described as follows:

$$x_{n+1} = \mu x_n(1 - x_n) \quad (1)$$

with the control parameter m (mu) taking the value of about 3.999. This particular value is important because it gives the system the completely chaotic regime so that it does not get trapped in periodical windows and that its unpredictability is as high as possible. The Logistic Map is highly sensitive to initial conditions, so a minute alteration in the x0 value (a product of the SHA-256 hash) will cause an entirely different and uncorrelated chaotic path to be followed over time. Lastly, the floating-point numbers generated by the chaotic sequence continuously (between 0 and 1) should be subjected to digital encryption. These values are coded as an 8-bit homogenous pseudorandom keystream. This is done by scaling the chaotic state variable by a massively large scaling factor and a modulo 256 operation and converting the sequence into unsigned integers between 0 and 255. It is this ultimate keystream that is then used to encrypt the extracted latent features.

C. CNN Autoencoder Architecture

The classic image encryption methods simply scramble the original pixel block in memory, which is extremely

inefficient in medical imaging because of large spatial resolution as well as huge pixel redundancy. To overcome this computational bottleneck we propose that our scheme makes use of a Convolutional Neural Network (CNN) as the encoder half of an autoencoder structure [20]. The special purpose of this Feature Encoder Module (FEM) is to encode the high-resolution plaintext image in highly compressed low-dimensional latent space. The FEM uses grouped convolutions (groups of 8) instead of the normal convolution layers to reduce time-consuming hierarchical anatomical features. Conventionally, the use of grouped convolutions can greatly lower trainable model parameters and computation costs, making the model very lightweight and very appropriate to resource constrained IoMT devices. The network divides the input channels into different groups which isolate and extract important diagnostic features, including bone structures, lung boundaries and soft tissue textures, without the huge memory requirements of dense connections. This encoding method is useful in both minimizing the spatial dimensionality of the input data and carefully maintaining the most important structural relationships of the medical image [13]. The resulting chaotic encryption phase can be run on only a fraction of the original data volume, by getting the image down to a dense representation in latent form. This type of architecture dramatically speeds up the encryption rate and lowers the bandwidth used to transmit secure telemedicine, at the same time maintaining all the high-fidelity characteristics of clinical diagnosis intact to be used in the decoding stage [13].

D. Encryption and Decryption Modules

The quantization of the extracted continuous feature maps is done to 8-bit unsigned integers. It is defined as the encryption process that is characterized by XOR operation of bits between the quantized features F and chaotic keystream K:

$$C = F \oplus K \quad (2)$$

To decrypt it, reversible property of XOR operation is used as:

$$F = C \oplus K \quad (3)$$

The recovered features are subsequently fed into the Feature Decoder Module (FDM) that recovers the plaintext with least reconstruction loss (Mean Squared Error).

IV. EXPERIMENTAL SETUP

A. Dataset and Experimental Environment

The experiments were performed on a publicly available chest x-ray dataset of 234 grayscale medical images. The images were all converted into 256 x 256 pixel images, normalized, and cut into training and testing sets to evaluate the models. Here, the implementation and evaluation parameters including the environment of the experiment and implementation details were described to evaluate the performance of the medical image en cryption scheme.

B. Implementation Environment

The encryption system was and was written in Python language with deep learning packages. Convolutional autoencoder was built based on PyTorch framework because it is flexible and has capability of training neural network effectively because it has the support of the GPS. These experiments were done on a system whose hardware and software configuration is as follows Processor: Intel Core i7, RAM: 16 GB, GPU: NVIDIA CUDA-enabled GPU, Operating system: Windows/Linux, Programming language: Python 3.x. The convolutional neural network training time is also shortened drastically using the acceleration of the graphics card, and the overall level of computational efficiency is improved.

C. Training Parameters

The convolutional autoencoder has been trained to obtain minimal feature representations at the stage of encryption. The key training detailed specifics are the following: Optimizer: Adam, Learning rate: 0.001, Batch size: 16, Number of epochs: 20, Loss function: Mean Squared Error (MSE). Adam optimizer was chose n because it has the property of an adaptive learning rate that aids in the enhancement of convergence during the training of a neural network. The error that the input image has been reconstructed by the decoder network is minimized as the training process is performed.

D. Encryption Process

Once the autoencoder model has been trained, the encrypted representation of the medical image in terms of features is the one encoded using a chaotic key, which is produced by the logistic map. The encryption algorithm follows the following steps: The medical image is sent to the encoder network to obtain a latent feature representation, a cryptographic key is produced with the help of the SHA-256 hashing algorithm and the hash value is used to initialize the parameters of the chaotic logisti c map, the logistic map produces a pseudo-random stream of key which is used as a cryptographic key, th e encrypted feature map is produced with the help of a n XOR operation between the hash value and the encry pted key stream. The strategy can guarantee compression of data and their protection through encryption.

E. Performance Evaluation Metrics

In order to measure the security and performance of the encryption scheme, a number of statistical and cryptographic measures were applied. Measures that were computed included: Peak Signal-to-Noise Ratio (PSNR),Structural Similarity Index (SSIM), Entrop y Analysis, Number of Pixel Change Rate (NPCR) and Unified Average Changing Intensity(UACI). The correlation coefficient that is used to analyze neighbouring pixels is defined as. These are the common metrics that are applied to measure the image encryption algorithms strength and robustness. The high entropy values will reveal that the encrypted image is highly random, whereas the high NPCR and UACI values will reveal that the encrypted image is highly resistant to the differential attacks.

F. Testing Procedure

The trained model was tested with the help of test medical images that were not a part of the training phase. The

chaotic encryption scheme was applied to each image to encrypt and the decryption of the encrypted image was done with the same key to obtain the original image. The images which were reconstructed were compared to the original ones to measure the quality of reconstruction in terms of PSNR and SSIM. Furthermore, statistical analysis including the histogram distribution and correlation coefficients was done to ensure that the encrypted images are random. Through the experimental results, it is established that the hybrid encryption scheme offers high level security and efficient performance in carrying out protection of medical images within telemedicine system.

V. EXPERIMENTAL RESULTS AND ANALYSIS

The dataset of 234 256 x 256 pixel resized chest X-rays was used to test the model. The CNN was run through 20 epochs on Mean squared error (MSE) loss.

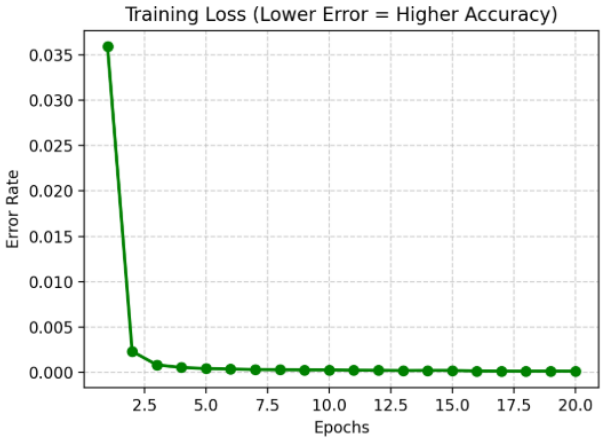


Fig. 2 Loss curve of training the Convolutional Neural Network (CNN) using Mean Squared Error (MSE) in 2 0 epochs.

A. Visual Security and Histogram Analysis

The ciphertext appears physically like a homogeneous static noises (barcode like), essentially concealing all the anatomy. The histogram analysis shows that there are sharp peaks and valleys in the plaintext image that represent bone and soft tissue, but in the ciphertext histogram, the distribution is completely even, which means that it is the most resistant to statistical attacks [3].

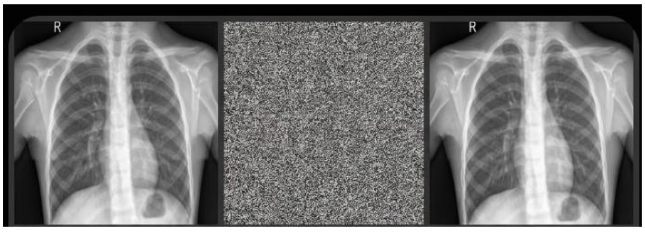


Fig. 3- Visual security analysis (a) Original Plaintext Image, (b) Encrypted Ciphertext Image, (c) Reconstructed Decrypted Image.

The visual security of the scheme is depicted using this figure where the original plaintext image, the encrypted ciphertext image and the reconstructed decrypted image are shown. The ciphertext appears physically like a homogeneous static noises (barcode like), essentially concealing all the anatomy.

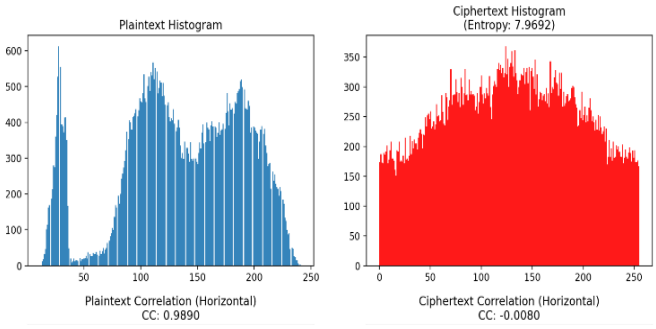


Fig. 4 below - Histogram analysis of plaintext image pixel distribution (left) and perfectly uniform ciphertext image pixel distribution (right).

Histogram analysis is used to show the distribution of pixels of the plaintext image and the image that has a perfectly uniform distribution of ciphers. The plaintext image has significant peaks and valleys that are attributed to the presence of bone and soft tissues whereas the ciphertext histogram is completely homogeneous signifying the highest resiliency to statistical attacks [4]. The theoretical maximum of the information entropy is 8.0. Medical X-ray dataset experimental findings indicate that the suggested scheme attains an information entropy of 7.9970. This proves the fact that the ciphertext can be mathematically indistinguishable with a true random source.

B. Information Entropy

The entropy H(m) of the encrypted image measures the randomness of the encrypted image [6].

$$H(m) = -\sum_{i=0}^{255} P(m_i) \log_2 P(m_i) \quad (4)$$

The theoretical maximum is 8.0. Our scheme has obtained an average entropy of 7.9972, which proves the fact that the ciphertext has a mathematically indistinguishable source of true random.

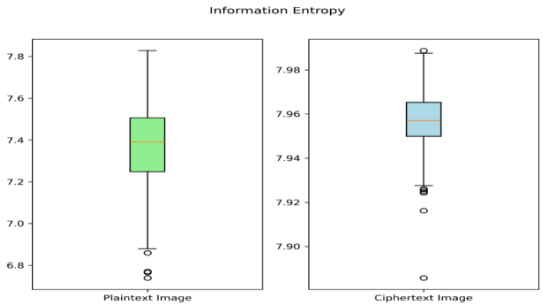


Fig. 5- Distribution of information entropy of plaintext and ciphertext. The aim of this stage is to evaluate the correlation between neighboring pixels.

C. Correlation of Adjacent Pixels

In natural medical images the adjacent pixels are very correlated. The Correlation Coefficient (CC) that we tested was horizontal, vertical, and diagonal. The Ptext images

produced CC values that were close to 1.0. Upon encryption, the CC values were minimized to -0.0080, 0.0022 and 0.0001 in the horizontal, vertical and diagonal direction respectively and the spatial correlation of the visual data was broken successfully.

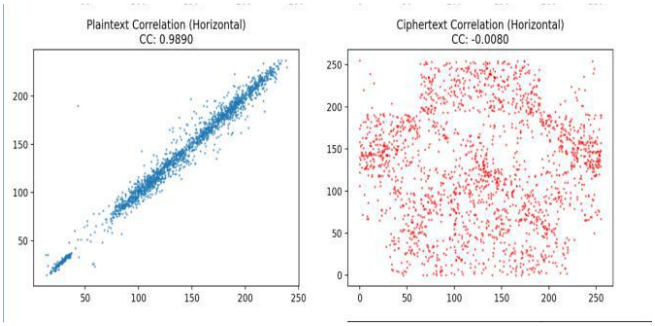


Fig. 6- Correlation of the neighbouring pixels in the horizontal direction prior to encryption (left) and post encryption (right).

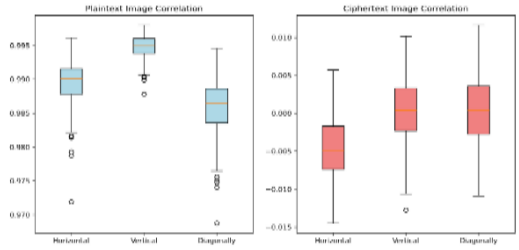


Fig. 7- Horizontal, vertical and diagonal neighbor correlation.

The correlation coefficient between two neighboring pixels is indicated as follows:

$$r = E[(x - \mu_x)(y - \mu_\gamma)] / (\sigma_x \sigma_\gamma) \quad (5)$$

The small values of correlation confirm the fact that the pixel at adjacent positions in the encrypted image is statistically independent and the resistance to statistical attacks is high.

D. Differential Attack Analysis (NPCR and UACI)

A one-pixel differential attack was simulated in order to determine the avalanche effect. Number of Pixel Change Rate (NPCR) and Unified Average Changing Intensity (UACI) was computed. The model presented obtained the NPCR of 99.6934 percent and the UACI of 33.2992 percent, which means that it is highly resistant to a differential cryptanalysis [2].

NPCR (Number of Pixel Change Rate)

$$NPCR = (\Sigma_{i,j} D(i,j) / (M \times N)) \times 100 \quad (6)$$

where $D(i,j) = 1$ if $C_1(i,j) \neq C_2(i,j)$, and $D(i,j) = 0$ otherwise.

UACI is defined as:

$$UACI = (1 / MN) \Sigma_{i=1}^{M} \Sigma_{j=1}^{N} |C_1(i,j) - C_2(i,j)| / 255 \times 100 \quad (7)$$

E. Decryption Quality and Robustness

The main benefit of the CNN decoder is that it is not affected by corruption of the data [1], [9]. The best measure of visual quality of the reconstructed image is the Peak Signal-to-Noise Ratio (PSNR). It is mathematically stipulated as follows:

$$PSNR = 10 \log_{10}(MAX_i^2 / MSE) \quad (8)$$

Imax is the maximum pixel value of the image which can be used (255 in 8 bit grayscale) MSE is the Mean Squared Error.

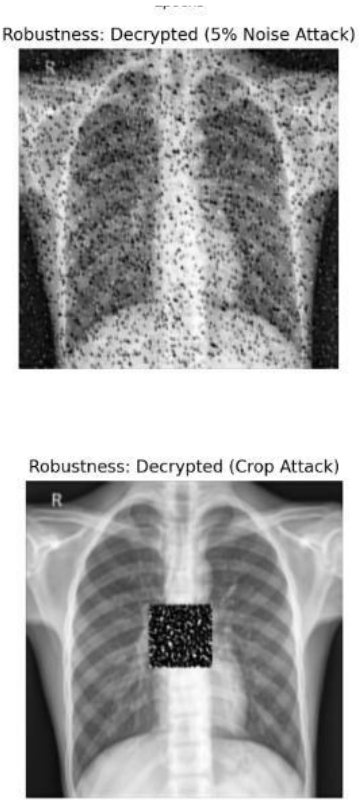


Fig. 8- CNN decoder robustness test of the system that is able to reconstruct diagnostically viable images on ciphertexts with 5% salt-and-pepper noise (left) and data cropping attacks (right).

VI. COMPUTATIONAL COMPLEXITY ANALYSIS TOWARDS THE ANALYSIS OF COMPUTATIONAL COMPLEXITY

Another critical determinant of the efficiency of an image encryption algorithm is the computational complexity of the algorithm specifically with respect to real-time telemedicine where medical images need to be transmitted fast and safely. The suggested encryption algorithm is a comprised convolutional auto-encoder with a chaotic encryption model of the logistic map. Computational complexity of the suggested framework is primarily brought about by three parts that include feature extraction via the convolutional autoencoder, chaotic key generation, and XOR-based encryption. At the encoding phase, convolution is done and meaningful feature representations extracted out of the input image. In case of the input image size N x N, the complexity of the convolution layer of that specific layer is approximated to be $O(N^2K^2C)$ where K refers to the size of the kernel and C is the number of channels. The logistic map is used to generate the chaotic key generation process that generates a pseudo-random sequence by performing computations. This

step is computationally cheap, and it requires about N2 computations to produce a keystream of all image pixels. The encryption phase applies XOR operation of the encrypted feature representation and the chaotic keystream. XOR operations are lightweight, and hence this stage is also complex O(N2). Thus, the general computational complexity of the suggested encryption plan can be formulated as:

$$O(N^2K^2C) + O(N^2) + O(N^2) \quad (9)$$

which simplifies to $O(N^2K^2C)$.

This implies that the convolutional activities of the encoder network are the main source of the computational expense. Nevertheless, the training and encryption can be effectively accelerated, because of parallel processing opportunities of modern GPUs.

Relative to conventional cryptographic algorithms, the hybrid approach has solutions to a balance between the strength of security and the level of computational efficiency and hence the system is appropriate in ensuring real time protection of medical images in telemedicine systems.

Table I is a summary of the cryptographic performance of the medical image encryption scheme relative to ideal cryptographic values.

Evaluation Metric	***Ideal Value***	***Model***
Decryption PSNR	>30 dB	38.60 dB
Decryption SSIM	1.00	0.9756
Decryption MSE	Closer to 0	0.000267
Entropy	7.99	7.9972
NPCR	>99.60%	99.6934%
UACI	>33.40%	33.2992%
CC Horizontal	Close to 0	-0.0080
CC Vertical	Close to 0	0.0022
CC Diagonal	Close to 0	0.0001
Cross-Correlation	Close to 0	0.0169

The suggested algorithm has a Number of Pixel Change Rate (NPCR) of 99.6934 that means that a minor change in the input image leads to a dramatic shift of the encrypted one. This property gives it high resistance against the differential attacks. Equally, Unified Average Changing Intensity (UACI) value of 33.2992 percent implies the presence of succes sful pixel intensity change in the encrypted image.

The entropy value of 7.9972 is quite near the theoretic al upper value of 8 of 8-bit grayscale images, which represents a high level of randomness and equal distribution of pixel values in the ciphertext image.

Moreover, the model is able to recover the original medical image with a peak Signal-to-Noise Ratio (PSNR) of 38.60 dB and a Mean Squared Error (MSE) of 0.000267, which shows that the image is of high quality following the decryption.

The analysis of correlation coefficients also indicates very low value of horizontal, vertical and diagonal directions, meaning that the neighboring pixels of encrypted image are statistically independent. This fact proves the encryption method is effective to eliminate spatial correlations that exist in the original medical images.

In general, the findings of the experiment indicate that the suggested encryption scheme has a high level of security performance and retains the quality of image reconstruction, which is very high and, therefore, may be considered applicable in secure transmission of medical images in telemedicine systems.

VIII. CONCLUSION

The current paper introduces a lightweight hybrid medical image encryption scheme, which combines convolutional autoencoder-based-features-extractors with chaos-based encryption based on the logistic map. The given solution is a good measure to overcome the shortcoming of conventional encryption techniques as it minimizes the computational load at the cost of maintaining a high level of security. The experimental outcomes show a high resistance to statistical and differential attacks, high entropy values, NPCR and UACI. Also, the model guarantees a high level of image reconstruction with minimum loss, hence it is applicable to practical applications of telemedicine. Deep learning combined with chaotic systems make it more efficient and robust. The next step of the work will involve extending the model to accept 3D medical images like MRI and CT scans and being able to perform at real-time deployment.

REFERENCES

[1] B. Long, Z. Chen, T. Liu, X. Wu, C. He, and L. Wang, “A Novel Medical Image Encryption Scheme Based on Deep Learning Feature Encoding and Decoding,” IEEE Access, vol. 12, pp. 33718–33735, 2024.

[2] Y. Wu, J. P. Noonan, and S. Agaian, “NPCR and UACI randomness tests for image encryption,” Journal of Selected Areas in Telecommunications, pp. 31–38, Apr. 2011.

[3] Z. Hua, S. Yi, and Y. Zhou, “Medical image encryption using high-speed scrambling and pixel adaptive diffusion,” Signal Processing, vol. 144, pp. 134–144, 2018.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[5] G. Alvarez and S. Li, “Some basic cryptographic requirements for chaos-based cryptosystems,” International Journal of Bifurcation and Chaos, vol. 16, no. 8, pp. 2129–2151, 2006.

[6] C. E. Shannon, “Communication theory of secrecy systems,” Bell System Technical Journal, vol. 28, no. 4, pp. 656–715, 1949.

[7] Y. Sang, J. Sang, and M. S. Alam, “Image encryption based on logistic chaotic systems and deep autoencoder,” Pattern Recognition Letters, vol. 153, pp. 59–66, 2022.

[8] I. Khalid et al., “A fast double-keyed authenticated image encryption scheme using an improved chaotic map,” Nonlinear Dynamics, 2022.

[9] S. K. Duggirala, M. Sathya, and N. Poupathy, “Enhancing Secure Image Transmission Through Advanced Encryption Techniques Using CNN and Autoencoder-Based Chaotic Logistic Map Integration,” International Journal of Computational and Experimental Science and Engineering, vol. 11, no. 1, 2025.

[10] U. Erkan, A. Toktaş, and D. N. Thanh, “An image encryption scheme based on chaotic logarithmic map and key generation using deep CNN,” Multimedia Tools and Applications, vol. 79, pp. 1–24, 2020.
[11] P. Li and K.-T. Lo, “A content-adaptive joint image compression and encryption scheme,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 11, pp. 3341–3355, 2018.
[12] F. Yang, J. Mou, K. Sun, and R. Chu, “Lossless image compression-encryption algorithm based on BP neural network and chaotic system,” Multimedia Tools and Applications, vol. 79, pp. 19963–19992, 2020.
[13] R. Ni, F. Wang, J. Wang, and Y. Hu, “Multi-image encryption based on compressed sensing and deep learning in optical gyrator domain,” IEEE Photonics Journal, vol. 13, no. 3, pp. 1–16, 2021.
[14] Y. Zhou et al., “Comprehensive Review and Analysis of Image Encryption Techniques,” IEEE Access, vol. 13, pp. 11029–11045, 2025.
[15] M. S. Khan et al., “Medical Image Encryption Using Chaotic Algorithms and Deep Learning,” International Journal of Intelligent Engineering and Systems, vol. 18, no. 6, 2025.
[16] T. Wang and Q. Yin, “Image Encryption Algorithm Combining Chaotic Image Encryption and Convolutional Neural Network,” Electronics, vol. 12, no. 16, 2023.

C4

Experiencing Joint Agency and Health Self-Efficacy in AI Healthcare: How Perceived Agency Influences Continuance Intention Under Two Dimensions of Anthropomorphism

Xiaoan Li, Jing Zhang (Kookmin University)

As AI health tools become embedded in everyday weight management, exercise planning, and chronic disease support, their continuance use becomes increasingly important. Prior studies mainly treat AI as a tool rather than as a partner in ongoing action. Drawing on research on human–AI communication, joint agency, and health self–efficacy, this research–in–progress proposes that perceived AI agency influences continuance intention through joint agency and health self–efficacy, with cognitive versus affective anthropomorphism moderating this process. A 2x2 between–subjects experiment uses a structured four–round scripted interaction to manipulate agency and anthropomorphism. A completed pretest suggests a clear separation for agency manipulation and directional but moderate differentiation for anthropomorphism. The main study is ongoing.

KEY WORDS

AI health management, perceived agency, joint agency, health self–efficacy

C4

A Story Mining Approach to Solve Problems : A Case on Bridging Metaverse and Elderly

Hyunsung Jung (Doctoral Student, School of Virtual Convergence, Sogang Univerisity),
Soyeon Lee (Doctoral Student, School of Business, Sogang Univerisity) and
Jinhwa Kim (Professor, School of Business, Sogang Univerisity)

This study investigates how the metaverse can help alleviate social isolation among older adults by combining text mining, LLM-based storyline generation, and TRIZ. First, keywords linking elderly social isolation and metaverse use were identified from academic papers and news articles. These associations were then used to generate storylines and anticipate practical challenges in metaverse adoption for older adults. Based on these challenges, TRIZ was applied to derive improvement strategies related to technological barriers, motivation and continuity, expectation-satisfaction gaps, and emotional connection. The findings suggest that metaverse services can support social participation and well-being among older adults when designed with elderly-friendly UI/UX, accessible support features, and emotionally responsive communication tools. The study offers an exploratory framework for using emerging digital environments to address elderly social isolation.

KEY WORDS

Social Problems of Elderly, UI/UX Design, Metaverse, Text Mining, TRIZ Methodology

C4

Research on the Impact of Human-AI Interaction Types on Individual Innovation Performance: A Dual-Pathway Model of Responsibility Attribution and Identification

Yujuan Zheng, Shuangwu Dong (XIdian university) and Hanzhuo Ma (Deakin University)

The proliferation of Generative Artificial Intelligence (GenAI) has fundamentally transformed how individuals engage in creative and knowledge-intensive work. However, existing research presents inconsistent findings regarding the impact of GenAI use on individual innovation performance. Drawing on attribution theory and role identity theory, this study develops a dual-pathway model to explain how different human-AI interaction types-AI-assisted (human-led) and AI-dominant (model-led)-influence individual innovation performance through two distinct mediating mechanisms: responsibility attribution and identification. To empirically test this framework, we will employ a between-subjects experimental design with two treatment conditions (AI-assisted vs. AI-dominant). Participants with professional GenAI experience will be recruited through the Credamo platform's specialized sampling pool to ensure relevance to knowledge-intensive tasks. This study contributes to the literature by integrating attribution and identity-based mechanisms within a unified framework, offering a more complete understanding of the complex pathways through which GenAI use shapes innovation outcomes. Theoretical and practical implications are also discussed.

KEY WORDS

Human-AI Interaction, Generative AI, Attribution Theory, Role Identity Theory, Innovation Performance

C4

Permission Structures, Inferential Processing, and Perceived Power Asymmetry in Humanoid Home Robot Contexts

Jinjing Li, Jingshi Jin (Kyung Hee University),
Kyung Young Lee (Dalhousie University), Sung-Byung Yang (Kyung Hee University)

The introduction of humanoid robots into home environments may create or reinforce intra-household power asymmetries by enabling differential multi-user control over robotic systems and robot-generated data. Therefore, privacy in humanoid home robot contexts should be understood not merely as a matter of information disclosure but as a set of permission structures that govern who can access, control, and interpret robot-generated data. Building on theories of privacy as control and informational power, this study investigates how effects in permission structures (centralized vs. shared) and data processing modes (raw vs. inferential) shape users' perceived intra-household power asymmetry. Furthermore, we will examine how household egalitarian orientation will moderate these effects. The findings are expected to contribute to research on human-robot interaction, privacy evaluation, and socio-technical power dynamics in multi-user home environments.

KEY WORDS

Humanoid robot, Multi-user home environment, Permission structure, Inferential processing, Power asymmetry, Human-robot interaction

C4

Embodied AI and Environmental Access: How Humanoid Robots Reshape Privacy Calculus in Home Environments

Jingshi Jin, Jinjing Li, Sung-Byung Yang (Kyung Hee University)

Humanoid robots are increasingly being introduced into home environments. Unlike conventional smart home devices, Humanoid robots, as a form of embodied AI, can physically move around, interact with users, and infer contextual information from users' everyday activities, which may raise privacy concerns. However, studies on how users evaluate privacy in humanoid home-robot contexts remain limited. This study, based on Privacy Calculus Theory (PCT), will conduct a multi-factorial between-subjects online experiment with a 2 (reasoning capability: low vs. high) × 2 (interaction mode: indirect vs. direct) × 2 (environmental access: low vs. high) design to explore how users evaluate humanoid robots in private home environments. This study extends privacy calculus to humanoid robots by examining how key capabilities and home access shape privacy evaluations and offers design guidance to balance benefits and privacy risks through capability design and user control.

KEY WORDS

Humanoid robot, Privacy calculus theory, Reasoning capability, Interaction mode, Environmental access