

# Generative AI in finance: Replicability, methodological contingencies, and future research directions

Hassnian Ali, Muhammad Bilal Zafar, Ahmet Faruk Aysan

## Item type

Journal Contribution

## Terms of use

This work is licensed under a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

## This version is available at

[https://manara.qnl.qa/articles/journal\\_contribution/Generative\\_AI\\_in\\_finance\\_Replicability\\_methodological\\_contingencies\\_and\\_fut](https://manara.qnl.qa/articles/journal_contribution/Generative_AI_in_finance_Replicability_methodological_contingencies_and_fut)

Access the item on Manara for more information about usage details and recommended citation.

Posted on Manara – Qatar Research Repository on

2025-11-04

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Finance Research Letters

journal homepage: [www.elsevier.com/locate/frl](http://www.elsevier.com/locate/frl)

## Generative AI in finance: Replicability, methodological contingencies, and future research directions

Hassnian Ali <sup>a,\*</sup> , Muhammad Bilal Zafar <sup>b,c</sup> , Ahmet Faruk Aysan <sup>d</sup> 

<sup>a</sup> Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

<sup>b</sup> Faculty of Social Sciences & Humanities, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

<sup>c</sup> School of Islamic Economics, Banking & Finance, Minhaj University Lahore, Lahore, Pakistan

<sup>d</sup> Hamad Bin Khalifa University, Qatar Foundation, Non-Resident Fellow Middle East Council on Global Affairs (ME Council), Research Associate, the University College London, Centre for Blockchain Technologies (UCL CBT), Qatar

### ARTICLE INFO

#### Keywords:

Generative AI  
Finance  
Bibliometric  
Systematic literature review  
Structural topic modeling

### ABSTRACT

Generative Artificial Intelligence (AI) is reshaping finance by transforming decision-making, risk management, and stakeholder engagement. This study provides a theory-informed synthesis of 84 peer-reviewed articles (2022–2025) using PRISMA-based screening, bibliometric analysis, and Structural Topic Modeling (STM). Six themes emerge: financial decision-making, ESG analytics, stock market prediction, advanced modeling for fraud detection and explainable AI, ChatGPT in accounting and education, and sentiment analysis with domain-specific LLMs. Findings show that generative AI enhances predictive capabilities and ESG assessments but raises issues of bias, transparency, and regulation. The review outlines future research priorities around interpretability, multimodal data, and governance frameworks.

### 1. Introduction

Generative Artificial Intelligence (AI) refers to computational systems capable of producing new content by learning from existing data (Feuerriegel et al., 2024). Its rapid diffusion into finance has redefined decision making, risk assessment, trading strategies, and customer engagement (Khan et al., 2025; Liu et al., 2024; Pelster and Val, 2024; Saxena et al., 2024; Schimanski et al., 2024; Wood et al., 2025). Large language models (LLMs) such as ChatGPT demonstrate unprecedented abilities in processing financial text, interpreting ESG disclosures, and automating client interactions (Fieberg et al., 2023; Jha et al., 2024; Jiang et al., 2022). Applications span portfolio optimization, credit scoring, sentiment analysis, and ESG evaluation, positioning generative AI as a transformative force across financial theory and practice (Y. Y. Cheng et al., 2024; Fatouros et al., 2024; Li et al., 2024).

Recent reviews establish a broad foundation but leave a clear gap for a generative AI-focused, theory-integrated synthesis. Bibliometric maps catalogue AI themes but treat the field as unified and predate the LLM inflection (Ahmed et al., 2022; Goodell et al., 2021). A bibliometric review of explainable AI refines interpretability debates yet overlooks assumptions and evaluation challenges posed by text-generative models (X.-Q. X.-Q. Chen et al., 2023). Generative-AI overviews broaden topics but omit combining systematic screening with quantitative structure discovery (Lee et al., 2024). Scoping work on ChatGPT in accounting and a consumer-behavior review probe subfield dynamics while leaving market microstructure, asset pricing, and ESG analytics underexamined (Dong et al., 2024; Meng et al., 2025). Methodological guidance on hybrid reviews advances synthesis but lacks an empirical

\* Corresponding author.

E-mail addresses: [Haal50943@hbku.edu.qa](mailto:Haal50943@hbku.edu.qa) (H. Ali), [bilalezafar@gmail.com](mailto:bilalezafar@gmail.com) (M.B. Zafar), [aaysan@hbku.edu.qa](mailto:aaysan@hbku.edu.qa) (A.F. Aysan).

<https://doi.org/10.1016/j.frl.2025.108797>

Received 27 August 2025; Received in revised form 7 October 2025; Accepted 29 October 2025

Available online 30 October 2025

1544-6123/© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

cross-thematic map (Pandey, 2025). The present study addresses these limits by centering explicitly on the post-2022 corpus, applying PRISMA screening, bibliometric mapping, and Structural Topic Modeling to produce a consolidated taxonomy with temporal dynamics, and linking those machine-discovered themes to finance theory while foregrounding replicability and methodological contingencies unique to generative models. The study is guided by three questions: (i) What are the publication trends, contributors, and leading journals in this field? (ii) Which thematic domains emerge from STM and what do they reveal about the scope of generative AI in finance? (iii) What methodological and theoretical gaps remain, and how do these shape future directions?

This review makes three contributions: (1) an empirically rigorous, PRISMA-filtered synthesis integrating bibliometric mapping

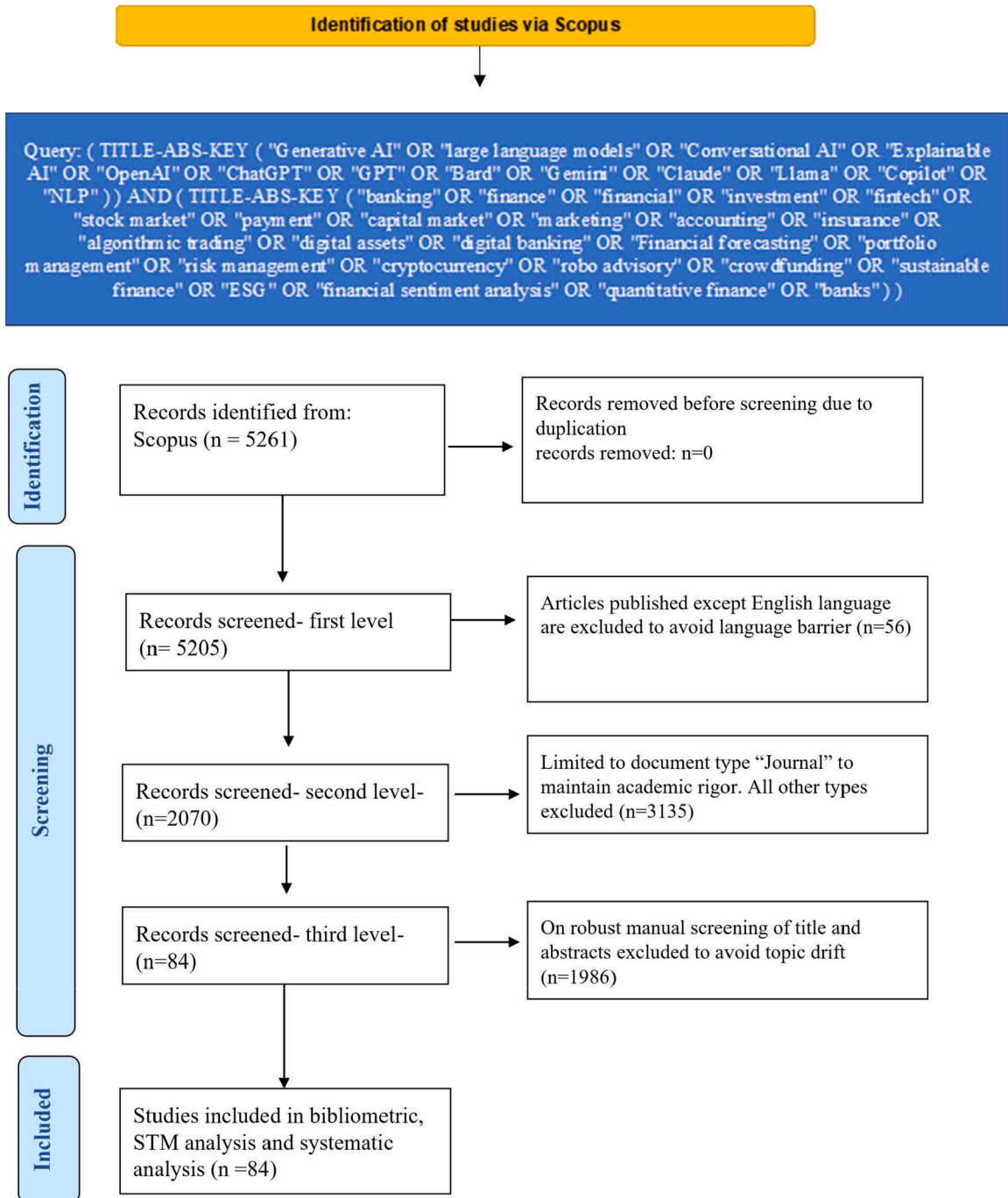


Fig. 1. PRISMA flow diagram. Adapted from (Page et al., 2021).

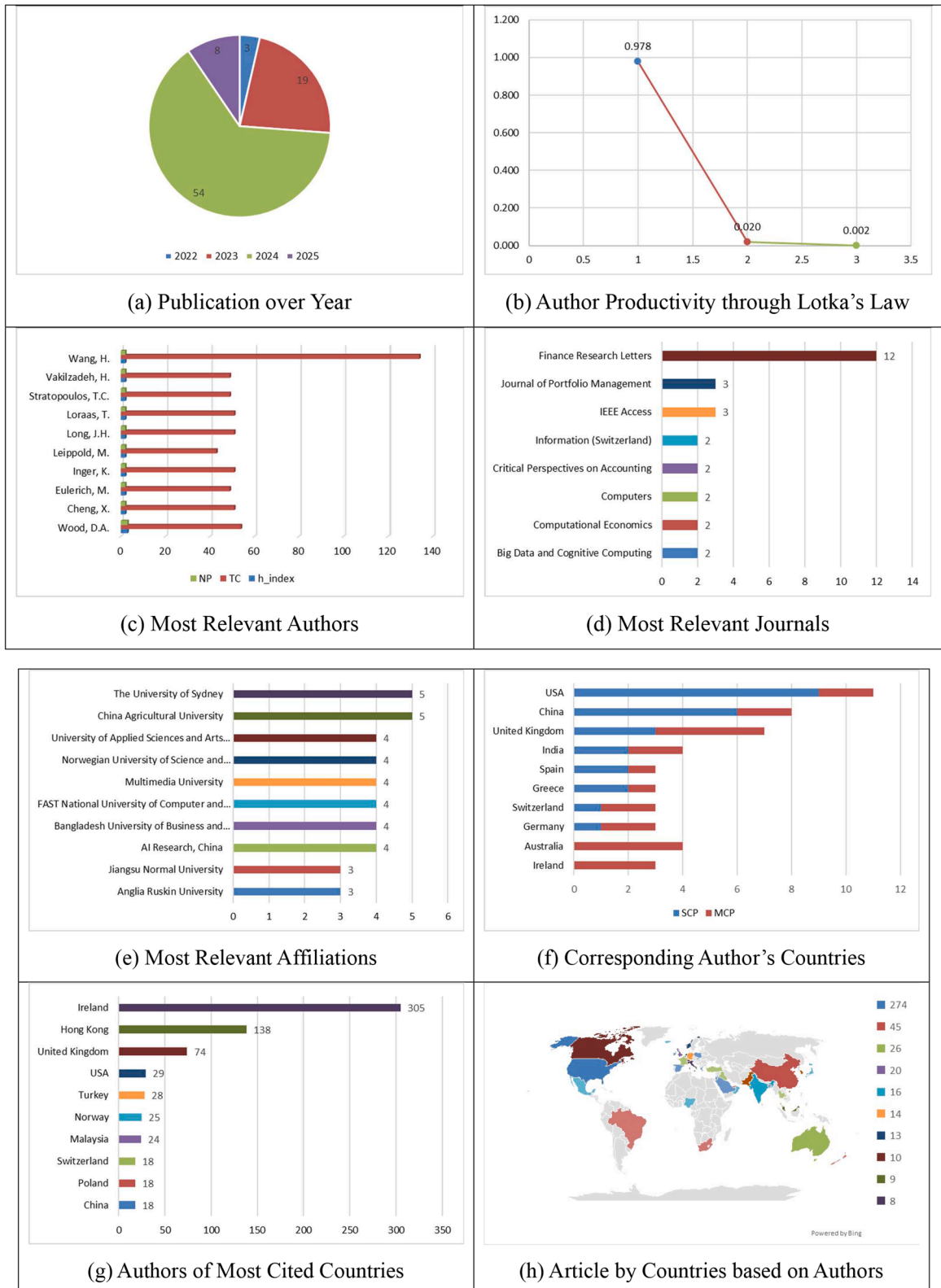


Fig. 2. Key Bibliometric Trends.

with STM to reveal temporal dynamics; (2) consolidation of established domains and identification of emergent trajectories (ESG LLMs, factor discovery, pedagogical shifts); (3) a focused agenda prioritizing replicability, theory-anchored evaluation, multimodal integration, and governance.

## 2. Data and methodology

### 2.1. Data collection and screening

The search was completed on 30 June 2025, and records from 2025 are therefore partial up to this date. The full Boolean query string and the inclusion–exclusion rules are shown in Fig. 1. Only peer-reviewed journal articles in English with explicit finance relevance were retained. Although this introduces some language-based bias, its primary implication is a potential underrepresentation of region-specific insights rather than a distortion of overall thematic patterns, given the global dominance of English in finance scholarship. Scopus was selected (Hussain et al., 2023; Muneeb et al., 2025; Pandey et al., 2024) for its breadth, replicability, and high overlap with Web of Science, which minimizes omission risk, though the absence of other databases is acknowledged as a limitation. PRISMA 2020 (Page et al., 2021) was selected to ensure replicability and transparency in screening, and the resulting 84 studies represent the most peer-reviewed and finance-relevant contributions in this emerging field.

### 2.2. Method

The study adopts a multi-method framework that integrates bibliometric analysis, Structural Topic Modeling (STM), and systematic review. Bibliometric analysis was conducted using Bibliometrix (Aria and Cuccurullo, 2017), mapping publication trends, journal concentration, collaboration patterns, and keyword co-occurrence to reveal the intellectual and structural contours of the field. STM was applied to the full text of all 84 articles, with publication year included as metadata to track topic prevalence over time (Roberts et al., 2019, 2014). STM was preferred over transformer-based approaches such as BERTopic for its replicable diagnostics and theoretical transparency. Each topic was then subjected to structured review, grounding findings in finance theory, synthesizing methods, and identifying empirical gaps. This integrative framework, summarized in Appendix Table A1, ensures that the analysis is both data-driven and theory-generative, enabling a rigorous synthesis of generative AI in finance.

## 3. Bibliometric findings

The bibliometric evidence highlights that between 2022 and 2025 (partial year 2025, up to 30 June), 84 articles on generative AI in finance were published across 58 journals, with an annual growth rate of 38.67 percent (Table 1). The field is young and relevant, with

**Table 1**  
Main information.

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2022:2025
Sources (Journals)	58
Articles	84
Annual Growth Rate %	38.67
Document Average Age	1.2
Average citations per doc	10.14
References	4112
DOCUMENT CONTENTS	
Keywords Plus (ID)	273
Author's Keywords (DE)	319
AUTHORS	
Authors	592
AUTHORS COLLABORATION	
Single-authored docs	9
Co-Authors per Doc	7.21
International co-authorships %	42.86

Publication activity is led by *Finance Research Letters*, followed by *IEEE Access* and the *Journal of Portfolio Management* (Fig. 2d), reflecting a growing dialogue between finance and computer science. Influential authors such as Wang, H. and Vakilzadeh, H. (Fig. 2c), alongside institutions like the University of Sydney, China Agricultural University, and FAST-NUCES (Fig. 2e), highlight concentrated productivity. The United States and China dominate in volume, while Ireland shows strong citation impact (Figs. 2f–2 g). The global map (Fig. 2h) illustrates broad diffusion, signaling that generative AI research is breaking disciplinary silos and fostering international collaboration.

an average article age of 1.2 years and more than ten citations per paper. Collaboration is extensive, involving 592 authors with over seven co-authors per study and 43 percent international contributions, showing rapid diffusion into an established domain.

### 3.1. Citation landscape and influential contributions

The citation analysis identifies several intellectual anchors. Dowling and Lucey's (2023) study on ChatGPT as a finance research tool leads the corpus with over 300 Scopus citations, highlighting both the potential and limitations of AI in knowledge generation (Dowling and Lucey, 2023). Huang et al. (2023) introduced FinBERT, a domain-specific LLM that improved sentiment and ESG-related text analysis. Wood et al. (2023) evaluated ChatGPT's performance in accounting education, revealing both utility and shortcomings in complex assessments. Collectively, these works illustrate a transition from early enthusiasm toward more critical evaluation of generative AI's epistemological limits (Leippold, 2023; Niszczota and Abbas, 2023). The most cited studies are summarized in Table 2, with topics ranging from portfolio optimization (Ko and Lee, 2024) to explainable credit assessment (Çelik et al., 2023; de Lange et al., 2022).

### 3.2. Collaboration and co-citation networks

Co-citation patterns (Appendix Fig. 1A) reveal clusters around authors such as Dowling, Huang, and Lopez-Lira, emphasizing sentiment analysis, ESG prediction, and financial risk modeling as central themes. Methodological alignment is evident, with clusters converging around tools such as FinBERT, GPT-3/4, and explainable AI frameworks. Collaboration networks (Appendix Fig. 2A) highlight prominent contributors, including Wood, Cheng, and Eulerich, who bridge institutions and disciplinary boundaries. These networks reflect how research is increasingly organized around consortia and interdisciplinary projects.

## 4. Topics identified through structural topic modeling

To move beyond descriptive bibliometric patterns and reduce subjective biases of manual coding, STM was applied to the corpus of 84 articles. STM estimates latent word distributions while incorporating document-level metadata, enabling analysis of both topic prevalence and variation over time (Roberts et al. 2014, 2019).

### 4.1. Model diagnostics

Diagnostic evaluation confirmed six coherent and interpretable topics. Fig. 3 reports held-out likelihood, residuals, semantic coherence, and lower bound across multiple values of  $K$ , tested with several random seeds to ensure stability. The choice of  $K = 6$  was guided by comparative diagnostics: while  $K = 5$  yielded slightly higher semantic coherence, it merged distinct conceptual domains (financial decision-making and ESG analytics), reducing interpretability;  $K = 7$ , in contrast, produced marginal gains in held-out likelihood but introduced topic fragmentation and higher residuals. Hence,  $K = 6$  offered the most balanced trade-off between model fit, coherence, and theoretical clarity (Rodriguez and Storer, 2020; Shin et al., 2023).

**Table 2**  
Most Cited Articles.

Paper Title	Authors	Total Citations (Scopus)	Total Citations (Google)	TC per Year (Scopus)	Normalized TC (Scopus)
ChatGPT for (Finance) research: The Bananarama Conjecture	(Dowling and Lucey, 2023)	302	568	100.6	8.589
FinBERT: A Large Language Model for Extracting Information from Financial Text*	(Huang et al., 2023)	132	393	44	3.754
The ChatGPT Artificial Intelligence Chatbot: How Well Does It Answer Accounting Assessment Questions?	(Wood et al., 2023)	46	107	15.3	1.308
From fiction to fact: the growing role of generative AI in business and finance	(B. X.-Q. Chen et al., 2023)	42	113	14	1.194
Sentiment spin: Attacking financial sentiment with GPT-3	(Leippold, 2023)	32	69	10.6	0.910
Extending machine learning prediction capabilities by explainable AI in financial time series prediction	(Çelik et al., 2023)	28	43	9.3	0.796
Bridging the gap—the impact of ChatGPT on financial research	(Cao and Zhai, 2023)	25	57	8.3	0.711
Explainable AI for Credit Assessment in Banks	(de Lange et al., 2022)	24	45	6	1.945
GPT has become financially literate: Insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice	(Niszczota and Abbas, 2023)	18	33	6	0.511
Can ChatGPT improve investment decisions? From a portfolio management perspective	(Ko and Lee, 2024)	14	58	7	5.142

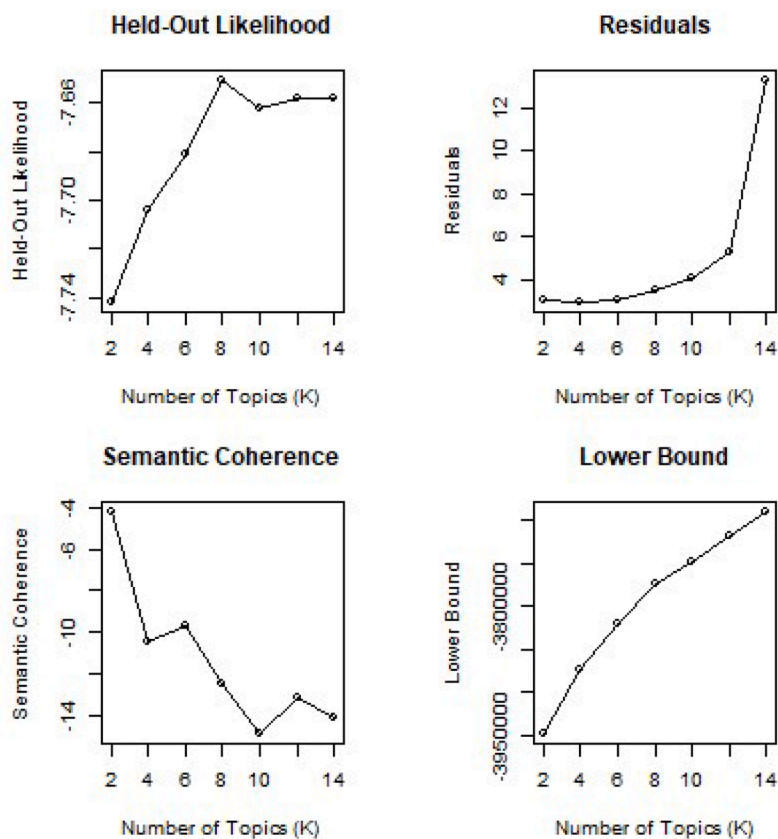


Fig. 3. Diagnostic tests for topic selection across values of K.

#### 4.2. Temporal dynamics

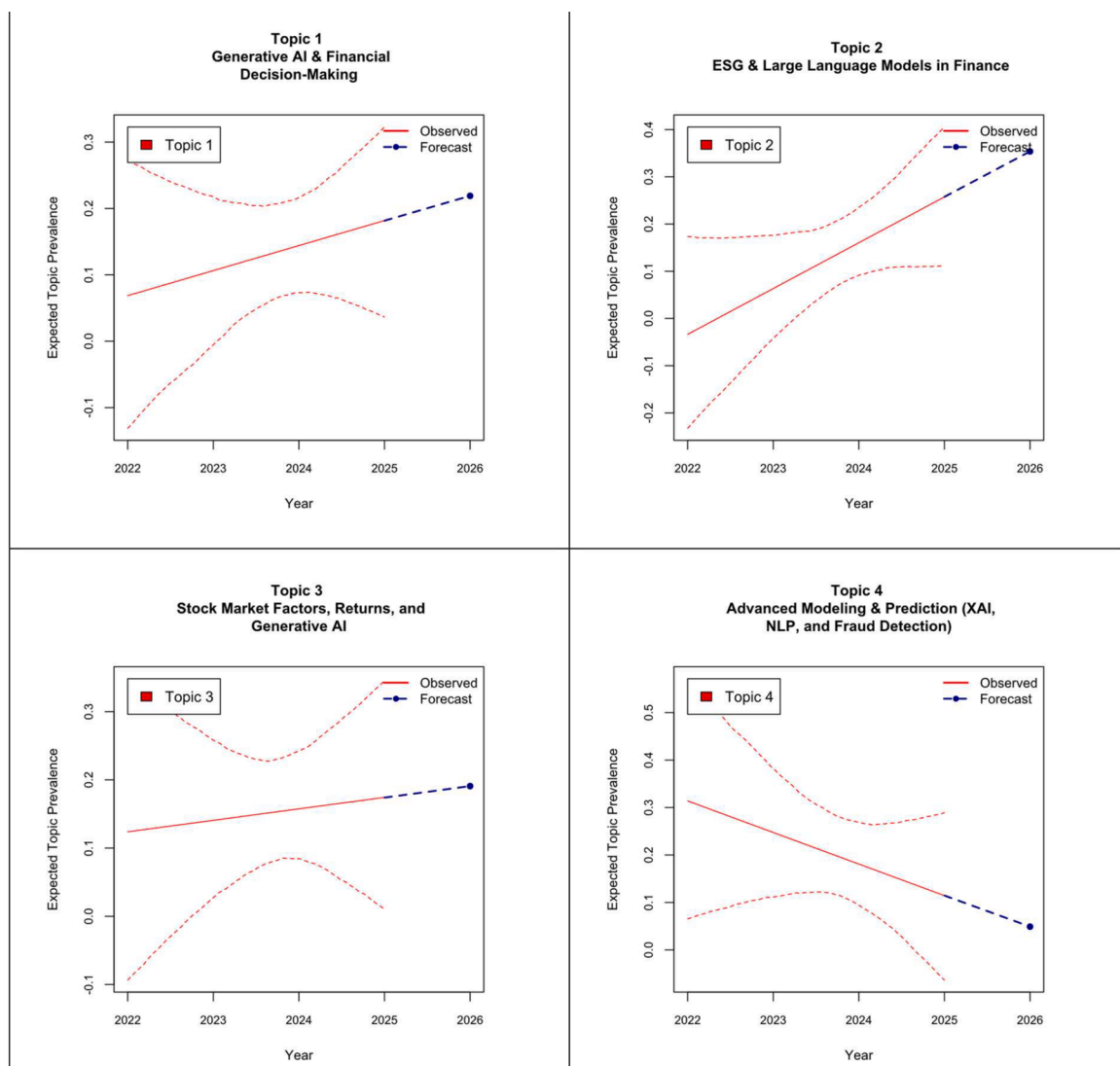
Fig. 4 illustrates topic prevalence between 2022 and 2025. Topics 1 (Generative AI in financial decision making), 2 (ESG and large language models), and 3 (Stock market factors and returns) display rising trajectories, reflecting growing scholarly attention to investment advice, sustainable finance, and predictive modeling. Topic 4 (Advanced modeling and prediction) has declined in relative prevalence, suggesting a shift away from technical modeling as a standalone focus. Topics 5 (ChatGPT in accounting and education) and 6 (Financial sentiment analysis) remain stable but specialized, indicating persistent niche interest. These temporal shifts highlight how research is evolving toward applied and integrative themes.

#### 4.3. Distinctiveness and interconnections

Figs. 5a–5c provide evidence of topic coherence, exclusivity, and network linkages. Topics 1 and 3 combine coherence with moderate exclusivity, indicating clearly defined yet connected domains. Topic 6 shows high exclusivity but weaker coherence, reflecting its focus on sentiment modeling with FinBERT and GPT variants. Correlation analysis (Fig. 5b) indicates generally weak associations among topics, with modest overlap linking ESG (Topic 2), returns (Topic 3), and sentiment (Topic 6). The network graph (Fig. 5c) corroborates this by showing sparse but non-random connections across clusters.

#### 4.4. Thematic taxonomy

Topic labels in Table 3 were human-assigned based on the most probable words and representative documents for each topic, following established STM interpretive protocols (Roberts et al., 2019). This ensured semantic precision and alignment with finance-specific conceptual categories. Topic 1, generative AI in financial decision making, shows how tools such as ChatGPT are applied to advisory functions, investment strategies, and financial literacy (Khan et al., 2025; Niszczota and Abbas, 2023). Topic 2, ESG and large language models, highlights the growing role of AI in analyzing corporate disclosures, measuring sustainability performance, and enabling real-time scoring systems (Balsiger et al., 2024; Wood et al., 2025). Topic 3, stock market factors and returns, captures how generative AI informs factor-based investing, volatility modeling, and return forecasting, particularly by integrating textual data into prediction (Ko and Lee, 2024; Wang, 2024). Topic 4, advanced modeling and prediction, covers explainable AI, credit risk, and fraud detection, where tools such as XAI improve transparency and predictive accuracy (Çelik et al., 2023; de Lange et al., 2022). Topic



**Fig. 4.** Topic Prevalence over Year (95 % Confidence Interval).

Note: The solid lines represent observed topic prevalence (2022–June 2025), while dashed segments denote partial-year and forecasted estimates beyond mid-2025.

5, ChatGPT in accounting and education, examines applications in reporting, training, and pedagogy, while raising concerns about academic integrity (Alshurafat et al., 2024; Wood et al., 2023). Topic 6, financial sentiment analysis, reflects strong interest in GPT and FinBERT for interpreting financial texts and linking sentiment to market outcomes (Huang et al., 2023; Pfeifer and Marohl, 2023).

## 5. Systematic review and implications

The six themes identified through STM provide a structured lens on how generative AI is reshaping finance while revealing common advances and constraints. Generative AI in financial decision-making positions tools such as ChatGPT as advisory aids and as experimental substitutes for human reasoning. Evidence links these models to portfolio choice, literacy tests, and investor profiling (Khan et al., 2025; Niszczota and Abbas, 2023; Ullah et al., 2024). The conceptual gain is scalable decision support. The persistent tension is performance that depends on trust, context, and user heterogeneity, which sustains demand for guardrails and evaluation protocols (Khan and Umer, 2024).

A second stream covers ESG and large language models. FinBERT and related systems outperform lexicons in detecting disclosure quality, greenwashing, and narrative integrity (Saxena et al., 2024; Schimanski et al., 2024; Wood et al., 2025). The conceptual shift is from static third-party ratings to dynamic text-driven scores that may realign legitimacy claims (Gao and Ju, 2024). Reliability is

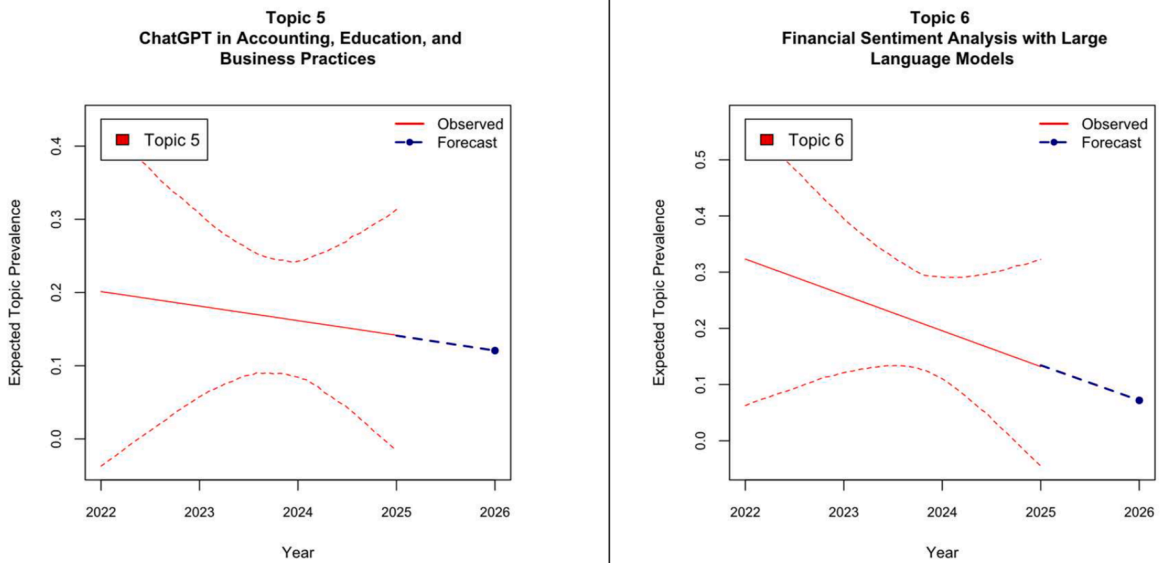


Fig. 4. (continued).

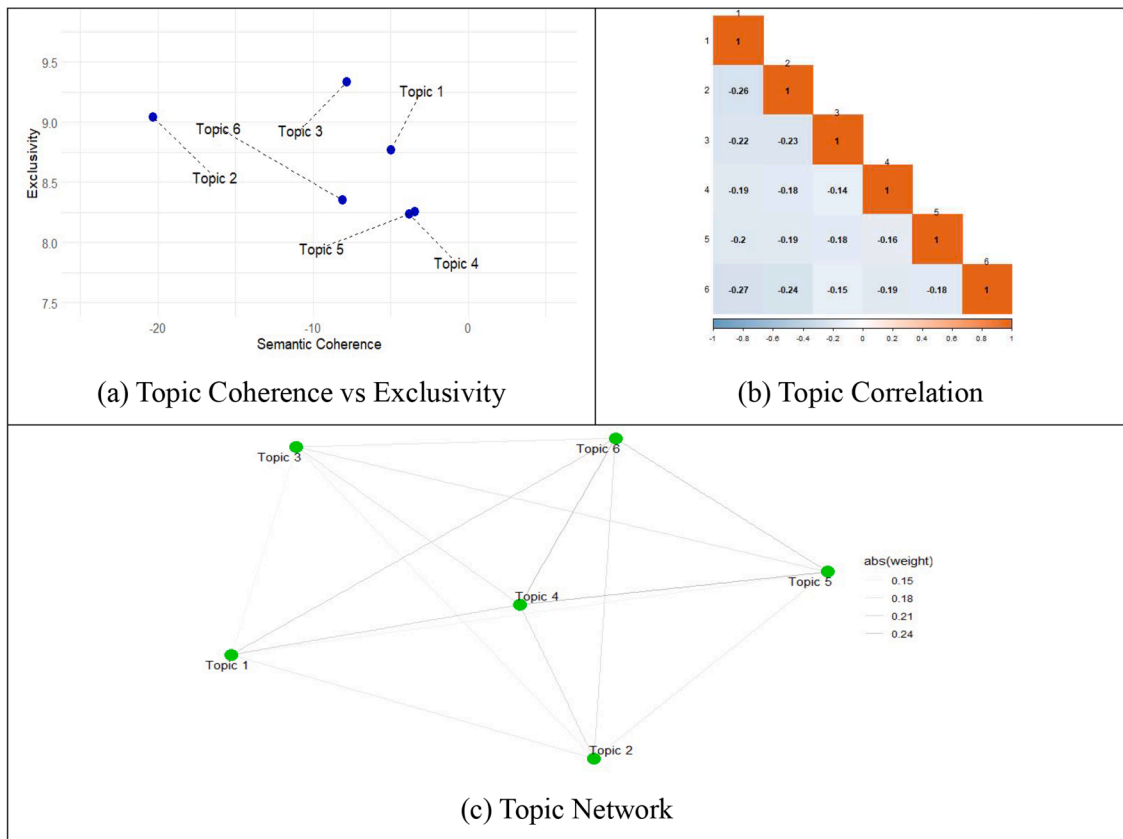


Fig. 5. Topic Coherence vs Exclusivity, Correlation and Network.

constrained by cross-cultural variation, strategic disclosure, and sparse data, which limit comparability and raise validity concerns (Wang, 2025).

The third theme on stock market factors and returns engages in asset pricing and market efficiency. Studies use generative models

**Table 3**  
Topics for Systemic Review based on STM.

Topic	Topic Prevalence	Topic Label	Label Words	Highest Prob Words	Label Rationale	Exemplary studies
Topic 1	0.21	Generative AI & Financial Decision-Making	financi, case, chatgpt, invest, decis	financi, case, chatgpt, invest, decis, make, bank, manag	These papers revolve around ChatGPT's role in financial advice, investment decisions, consumer decision-making processes, and personal finance issues.	(Khan et al., 2025; Moharrak and Mogaji, 2024; Niszczota and Abbas, 2023; Schlosky et al., 2024; Trozze et al., 2024)
Topic 2	0.21	ESG & Large Language Models in Finance	esg, model, llms, score	esg, model, llms, score, invest, question, financi, llm	Focuses on environmental, social, and governance (ESG) analytics, discussing how large language models can assess ESG commitments, measure ESG performance, or interpret ESG disclosures.	(Balsiger et al., 2024; Fairhurst and Greene, 2024; Gao and Ju, 2024; Wood et al., 2025)
Topic 3	0.18	Stock Market Factors, Returns, and Generative AI	factor, financi, stock, return, market	factor, financi, stock, return, market, chatgpt, portfolio, model, risk, data	Centers on portfolio returns, factor modeling, herding behavior in generative AI stocks, and how ChatGPT can enhance or forecast stock returns.	(Y. Y. Cheng et al., 2024; Fior et al., 2022; Ko and Lee, 2024; Li et al., 2024; Wang, 2024)
Topic 4	0.15	Advanced Modeling & Prediction (XAI, NLP, and Fraud Detection)	model, predict, data, featur	model, predict, data, featur, valu, financi, learn, market	Encompasses explainable AI (XAI) and next-generation modeling approaches for credit scoring, fraud detection, or time-series predictions, focusing on how these data-driven methods are applied in finance.	(Çelik et al., 2023; de Lange et al., 2022; Fior et al., 2022; Nobel et al., 2024; Thanathamathee et al., 2024)
Topic 5	0.13	ChatGPT in Accounting, Education, and Business Practices	account, chatgpt, use, model	account, chatgpt, model, univ, technolog, busi, generat, questio	Emphasizes ChatGPT's usage and implications for accounting tasks, education, and business frameworks (e.g., exam performance, academic dishonesty, professional transformations).	(Alshurafat et al., 2024; Haugland Sundkvist and Kulset, 2024; Muzanenhano and Power, 2024)
Topic 6	0.13	Financial Sentiment Analysis with Large Language Models	model, sentiment, financi, data	model, sentiment, financi, data, text, languag, dataset, learn, news	Captures domain-specific LLMs (e.g., FinBERT, GPT-based methods) for sentiment classification in financial text, from news and earnings calls to social media, focusing on how that sentiment influences market behavior.	(Delgadillo et al., 2024; Dogra et al., 2022; Huang et al., 2023; Lan et al., 2024; Pfeifer and Marohl, 2023)

to mine textual disclosures and construct signals that correlate with abnormal returns and volatility (Y. Y. Cheng et al., 2024; Fatouros et al., 2024; Li et al., 2024). The conceptual advance is factor discovery from unstructured text. Constraints arise in non-English settings, small-cap coverage, and live trading confirmation, which limit external validity and theory integration (Fatouros et al., 2024).

Advanced modeling and prediction form a fourth cluster that spans fraud detection, credit risk, and forecasting. Work with XGBoost, LSTM, and FinChain BERT exceeds benchmark accuracy and uses SHAP and LIME to increase transparency for compliance needs (de Lange et al., 2022; Nobel et al., 2024). The conceptual contribution is a workflow that pairs high accuracy with post hoc explanation. The limiting factor is ecological validity since synthetic data, such as PaySim reduces replicability in real banking contexts (Awosika et al., 2024).

A fifth theme examines ChatGPT in accounting, education, and business practice. Tools improve summarization and study support, but underperform in technical accounting reasoning and can weaken academic integrity, where the fraud triangle theory helps explain misuse of incentives (Alshurafat et al., 2024; Muzanenhano and Power, 2024; Wood et al., 2023; Zhao and Wang, 2024). The conceptual point is augmentation rather than substitution. The constraint is accountability and epistemic justice in varied institutional settings, including non-Western contexts (Eulerich et al., 2024).

Finally, financial sentiment analysis remains mature yet evolving. FinBERT, GPT, and LLaMA families beat lexicons across earnings calls, bonds, and crypto, which institutionalizes text as a predictive input (Chiu and Hung, 2025; Huang et al., 2023; Roumeliotis et al., 2024). The conceptual gain is a stable pipeline that links narrative tone to market outcomes. Open problems include sarcasm, figurative language, and multilingual corpora, which motivate ensemble designs and real-time dashboards with interpretable metrics (Delgadillo et al., 2024; Lan et al., 2024; Mathebula et al., 2024).

Taken together, the six themes converge on three cross-cutting insights. First, generative AI yields measurable predictive and process gains, yet these gains are conditional on setting, data provenance, and user trust. Second, the field is moving from static artifacts to dynamic, text-informed signals, which reframes legitimacy in ESG, signal construction in asset pricing, and assurance in accounting. Third, explanation and governance remain central, which underscores the need for a technology-ethics-regulation lens where performance is assessed alongside regulatory criteria such as the EU AI Act's (Quintais, 2025) high-risk system classification and China's Interim Measures for Generative AI Services (Migliorini, 2024).

To guide future work, three priorities emerge that bridge theory and practice. First, generalizability across languages and markets is essential for credible inference and for adoption by global institutions. Second, prospective and live trading validation is needed to connect backtests with market efficiency debates and to calibrate risk. Third, governance for high-stakes use cases such as ESG scoring and fraud detection should integrate explainability with audit trails and regulatory comparators within this technology–ethics–regulation framework.

Table 4 organizes the broader gap landscape and aligns it with actionable directions.

**Table 4**  
Limitations and Future Research Gaps in Generative AI & Finance.

Limitations and Gaps	Authors	Future Research Directions
<b>Data and Model Limitations</b>		
Models struggle to incorporate and generalize cross-market and multilingual financial data.	(Lan et al., 2024; Mathebula et al., 2024; Schimanski et al., 2024)	Develop multilingual and cross-market datasets to enhance LLM applicability in diverse financial contexts.
LLMs depend heavily on publicly available text, which may be biased, misleading, or superficial, especially in ESG analysis.	(Schimanski et al., 2024; Wang, 2025; Wood et al., 2025)	Enhance data preprocessing techniques and validate textual data sources for more reliable AI-driven ESG insights.
Lack of domain-specific datasets and fine-grained data for sentiment and ESG analysis, particularly in niche languages like Chinese.	(Dogra et al., 2022; Lan et al., 2024)	Create specialized datasets with expert annotations for financial domains and underrepresented languages.
Social media data poses challenges due to its noise, lack of accuracy, and susceptibility to manipulation.	(Mathebula et al., 2024; Roumeliotis et al., 2024)	Develop robust preprocessing and filtering techniques to address the inherent limitations of social media data.
<b>Methodological Limitations</b>		
Models often struggle with sarcasm, irony, and figurative language in financial texts, leading to inaccuracies in sentiment analysis.	(Dogra et al., 2022; Faccia et al., 2024; Mathebula et al., 2024)	Improve aspect-based sentiment analysis and develop models trained specifically on figurative and domain-specific language.
Limited research evaluates the accuracy of LLMs in specific market segments like small-cap or emerging markets.	(Y. Y. Cheng et al., 2024; Fatouros et al., 2024)	Conduct targeted studies to assess LLM performance in underrepresented market segments.
The lack of real-time data integration limits the use of LLMs for dynamic financial decision-making and sentiment analysis.	(Delgadillo et al., 2024; Mathebula et al., 2024)	Develop hybrid real-time LLM systems to provide immediate insights for decision-making.
Limited empirical validation of LLM-based ESG scoring in comparison to traditional ESG metrics from agencies like MSCI and Sustainalytics.	(Schimanski et al., 2024; Wood et al., 2025)	Conduct comparative studies to benchmark AI-driven ESG scores against traditional methodologies.
<b>Ethical Concerns</b>		
Greenwashing detection remains a critical challenge, as LLMs rely solely on textual data and fail to capture inconsistencies between disclosures and practices.	(Schimanski et al., 2024; Wang, 2025)	Incorporate multimodal data (text, images, and videos) to improve the detection of greenwashing in ESG reports.
The potential misuse of generative AI for market manipulation remains unaddressed in many studies.	(Khan and Umer, 2024; Zheng et al., 2024)	Research regulatory frameworks and monitoring tools to identify and mitigate manipulative practices using generative AI.
<b>Performance Issues</b>		
LLMs struggle to process and analyze long and complex financial documents such as 10-K filings and earnings conference transcripts.	(Chiu and Hung, 2025; Dogra et al., 2022)	Implement document summarization techniques using advanced NLP models like LongT5 to improve processing efficiency.
Performance of LLMs in live trading scenarios remains underexplored, with most studies focusing on backtesting.	(Fatouros et al., 2024; Ko and Lee, 2024)	Conduct real-world evaluations of AI-driven investment strategies in live trading environments.
Current sentiment analysis models lack robustness in assessing financial sentiment across different asset classes, including crypto and equities.	(Huang et al., 2023; Roumeliotis et al., 2024)	Enhance sentiment analysis models to handle domain-specific nuances and varying market conditions.
<b>Application-Specific Challenges</b>		
Limited integration of LLMs with financial workflow tools and accounting systems.	(Li and Vasarhelyi, 2024; Zhao and Wang, 2024)	Research the seamless integration of LLMs into existing financial and accounting software ecosystems.
ChatGPT and similar models struggle with accounting-specific tasks like journal entries and financial statement preparation.	(X. Y. Cheng et al., 2024)	Fine-tune generative AI models for accounting tasks and integrate them with accounting standards like IFRS and GAAP.
The absence of real-world case studies evaluating the impact of generative AI in advanced modeling tasks, such as fraud detection.	(Nobel et al., 2024; Thanathamathhee et al., 2024)	Develop real-world case studies focusing on fraud detection and financial risk assessment using generative AI.
<b>Theoretical Gaps</b>		
The impact of generative AI on financial literacy and inclusion in developing economies remains largely unexplored.	(Khan et al., 2025; Liu et al., 2024)	Conduct interdisciplinary research to understand the role of AI in improving financial literacy and access in underprivileged regions.
Limited focus on integrating AI-driven insights with traditional financial theories such as CAPM and behavioral finance.	(Y. Y. Cheng et al., 2024; Li et al., 2024)	Propose hybrid theoretical models that combine AI-driven insights with established financial frameworks.
Insufficient exploration of how human-AI collaboration can improve financial decision-making.	(Khan and Umer, 2024; Ullah et al., 2024)	Develop theoretical frameworks and pilot studies to evaluate collaborative financial decision-making between humans and AI systems.

## 6. Conclusion and limitations

This study provides a systematic and theory-informed review of generative AI in finance, synthesizing evidence from 84 peer-reviewed articles through bibliometric mapping, STM, and structured thematic analysis. By identifying six interrelated themes, it demonstrates how generative AI extends decision theory, stakeholder theory, market efficiency, and explainable AI ethics while also exposing tensions between predictive gains and replicability, innovation and ecological validity, and efficiency and governance. The review is not without limitations. The dataset is restricted to a relatively small sample of journal articles due to the infancy stage of the area of investigation. Scopus was selected given its breadth, replicability, and high overlap with Web of Science (WoS), and therefore WoS was not separately added. The analysis is also limited to English-language publications, which may underrepresent non-English scholarship, and to academic corpus data without expert interviews or surveys. Despite these constraints, the study advances a coherent conceptual framework for understanding generative AI in finance.

### Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author(s) used *ChatGPT-5* in order to refine sentences and improve clarity. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the final content of the manuscript.

### CRediT authorship contribution statement

**Hassnian Ali:** Writing – review & editing, Writing – original draft, Resources, Methodology, Data curation, Conceptualization. **Muhammad Bilal Zafar:** Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Conceptualization. **Ahmet Faruk Aysan:** Writing – review & editing, Validation, Supervision, Methodology, Data curation, Conceptualization.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.frl.2025.108797](https://doi.org/10.1016/j.frl.2025.108797).

## Appendix

[Table A1](#), [Fig. A1](#), [Fig. A2](#)

**Table A1**  
Comprehensive systematic review framework.

Category	Topic 1: Generative AI & Financial Decision-Making	Topic 2: ESG & Large Language Models in Finance	Topic 3: Stock Market Factors, Returns, and Generative AI	Topic 4: Advanced Modeling & Prediction (XAI, NLP, and Fraud Detection)	Topic 5: ChatGPT in Accounting, Education, and Business Practices	Topic 6: Financial Sentiment Analysis with Large Language Models
<b>1. Scope &amp; Conceptualization</b>	How is generative AI (e.g., ChatGPT) conceptually defined and applied within financial decision-making contexts (e.g., portfolio selection, credit decisions, personal finance), and which theoretical lenses (e.g., decision theory, behavioral finance) are predominantly used?	How are large language models (e.g., ChatGPT, FinBERT) being used to assess or score ESG disclosures and commitments, and which types of ESG data (annual reports, corporate social responsibility statements) are commonly analyzed?	How do authors frame generative AI in the context of traditional asset pricing theories or factor models (CAPM, Fama-French, etc.), and are new theoretical constructs proposed for AI-driven factor discovery?	Which machine learning or deep learning techniques (XGBoost, BERT, GPT-based, LSTM) are most common for prediction tasks in finance (fraud detection, credit scoring, etc.), and how do these models handle imbalanced data often found in fraud contexts?	How is ChatGPT being adopted in accounting tasks (bookkeeping, auditing, reporting), and what are student/instructor perceptions of ChatGPT for educational use (e.g., exam prep, assignment assistance)?	What LLM-based methods (FinBERT, GPT variants) are used for financial sentiment, and how do they compare to older lexical or machine learning approaches in terms of accuracy and contextual understanding?
<b>2. Methodological Approaches</b>	What empirical methods do authors use to examine the effectiveness of AI-	What methodological frameworks (lexicon-based vs. machine learning)	What AI-driven models (GPT variants, transformers) are used to predict	How are XAI techniques (SHAP, LIME, Anchors) implemented to	Which accounting subfields (financial accounting,	Which textual data sources—social media, news articles, earnings calls—dominate

(continued on next page)

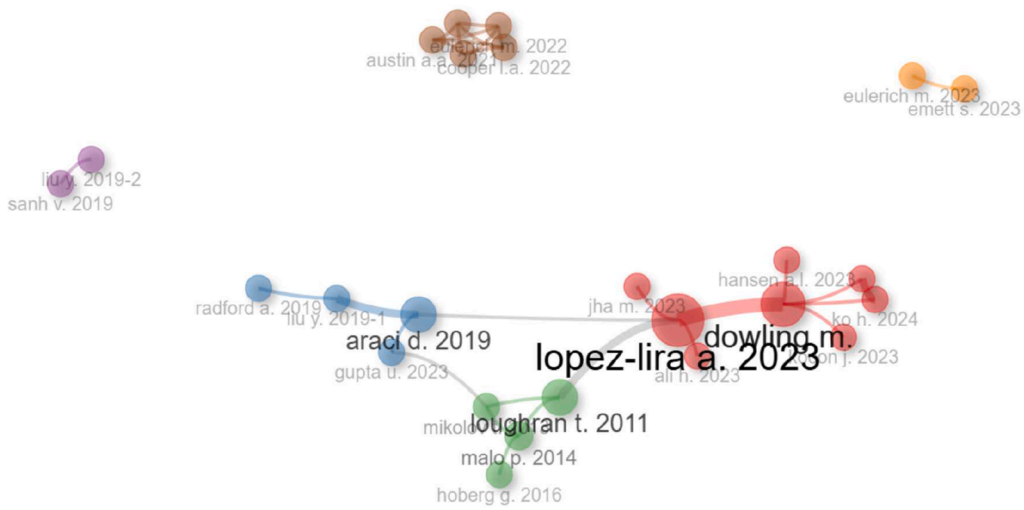
Table A1 (continued)

Category	Topic 1: Generative AI & Financial Decision-Making	Topic 2: ESG & Large Language Models in Finance	Topic 3: Stock Market Factors, Returns, and Generative AI	Topic 4: Advanced Modeling & Prediction (XAI, NLP, and Fraud Detection)	Topic 5: ChatGPT in Accounting, Education, and Business Practices	Topic 6: Financial Sentiment Analysis with Large Language Models
	driven or ChatGPT-driven decision-making (e.g., surveys, lab/field experiments, case studies), and how do studies measure improvements (or lack thereof) in decision quality?	dominate for ESG text analysis, how are they validated, and how do LLM-driven ESG scores compare with traditional ESG rating agencies (e.g., MSCI, Sustainalytics)?	stock returns, how do they compare against standard factor models, and which metrics (RMSE, MAPE, alpha generation) are used to benchmark performance?	interpret predictions in high-stakes financial tasks, and are transparency and trust addressed, especially regarding regulatory compliance or stakeholder acceptance?	management accounting, auditing) are seeing more AI integration, and are there distinct differences in usage between educational vs. professional settings?	research on financial sentiment, and are multi-lingual or cross-market studies common?
<b>3. Outcomes &amp; Impact</b>	What performance metrics (e.g., decision accuracy, cost reduction, user satisfaction) are used to assess generative AI's role in financial decisions, and are there contexts (e.g., retail vs. institutional investors, developed vs. emerging markets) where AI-based support shows stronger or weaker effects?	How does improved ESG measurement via LLMs translate into investment decision-making or risk management, and are there studies linking LLM-generated ESG scores to firm performance, cost of capital, or reputational risk?	Do LLM-derived signals (e.g., news sentiment, ChatGPT analysis) imply market inefficiencies or behavioral biases, and are certain market segments (e.g., small-cap, emerging markets, crypto) more sensitive to AI-driven factor discoveries?	What metrics do authors use to evaluate predictive models (accuracy, AUC, recall for fraud detection), and are there recognized public benchmarks/datasets, or do authors rely heavily on proprietary data?	Does ChatGPT usage enhance or undermine critical thinking and ethical decision-making for accounting students, and which frameworks or guidelines do authors propose to mitigate academic dishonesty?	How do sentiment signals derived from LLMs correlate with trading volume, asset volatility, or price returns, and are certain asset classes (crypto vs. equities) more reactive to sentiment changes?
<b>4. Ethical &amp; Regulatory Considerations</b>	How do authors address bias, data privacy, or compliance issues in AI-driven decisions, and what regulatory frameworks are proposed or needed for the adoption of generative AI in high-stakes finance?	What are the main limitations (e.g., data availability, domain-specific jargon, "greenwashing" detection) in applying LLMs to ESG, and how do authors handle cross-cultural or language differences in ESG reporting?	How do authors address the idea of epistemic injustice, data biases, or misinformation from AI in accounting contexts, and are there frameworks for ensuring transparency when ChatGPT is used in producing or explaining financial statements?	Are researchers applying XAI or adversarial testing to ensure sentiment models are robust and less prone to manipulation, and how do they address issues like sarcasm, domain-specific jargon, or "noise" in financial text?		
<b>5. Research Gaps &amp; Future Directions</b>	Where are the major research gaps (e.g., few real-world experiments, limited user acceptance data), and which open questions remain regarding interpretability or user trust in AI-based financial advisement?	Which emerging techniques (multimodal data, advanced XAI for ESG transparency) are suggested, and are there calls for new regulatory guidelines on AI-based ESG ratings?	What are the data limitations and model interpretability challenges, and how do studies suggest we integrate non-traditional data sources (e.g., social media, climate data) into factor models using AI?	What new frontiers (e.g., unsupervised anomaly detection, Federated Learning) might redefine the scope of advanced modeling in finance, and are there calls to standardize or unify approaches to XAI in	What do studies suggest regarding curriculum design or professional training to accommodate ChatGPT-like tools, and are there calls for policy or regulatory changes in how AI outputs are integrated or audited?	Which emerging directions—e.g., real-time sentiment tracking, ensemble approaches, multimodal data—are likely to shape next-generation sentiment analysis, and is there an identified need for standardized financial sentiment benchmarks?

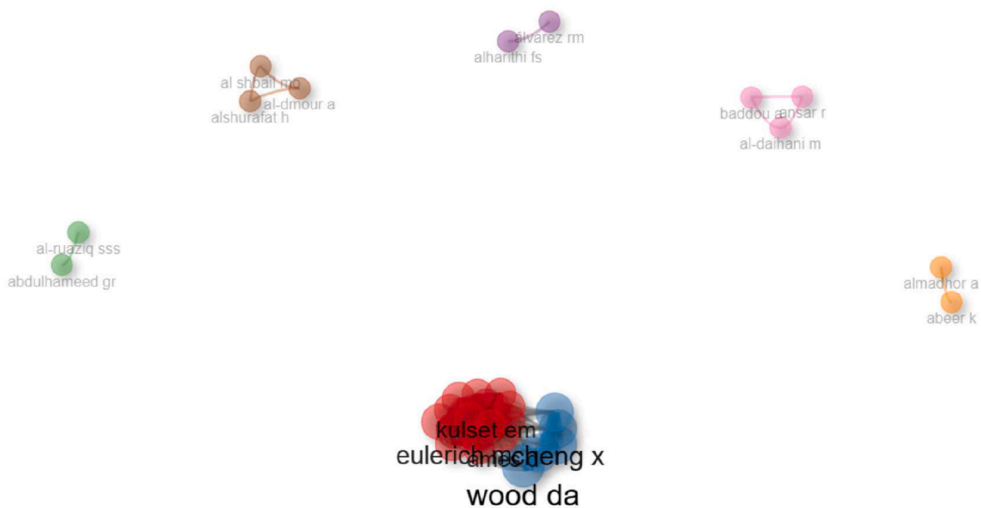
(continued on next page)

**Table A1** (continued)

Category	Topic 1: Generative AI & Financial Decision-Making	Topic 2: ESG & Large Language Models in Finance	Topic 3: Stock Market Factors, Returns, and Generative AI	Topic 4: Advanced Modeling & Prediction (XAI, NLP, and Fraud Detection)	Topic 5: ChatGPT in Accounting, Education, and Business Practices	Topic 6: Financial Sentiment Analysis with Large Language Models
				financial contexts?		



**Fig. 1A.** Co-Citation Networks.



**Fig. 2A.** Collaboration Networks.

**Data availability**

I have uploaded files as supplementary material

**References**

Ahmed, S., Alshater, M.M., Ammari, A.El, Hammami, H., 2022. Artificial intelligence and machine learning in finance: a bibliometric review. Res. Int. Bus. Financ. 61, 101646. <https://doi.org/10.1016/j.ribaf.2022.101646>.

- Alshurafat, H., Al Shbail, M.O., Hamdan, A., Al-Dmour, A., Ensour, W., 2024. Factors affecting accounting students' misuse of chatgpt: an application of the fraud triangle theory. *J. Financ. Report. Account.* 22, 274–288. <https://doi.org/10.1108/JFRA-04-2023-0182>.
- Aria, M., Cuccurullo, C., 2017. bibliometrix: an R-tool for comprehensive science mapping analysis. *J. Informetr.* 11, 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>.
- Awosika, T., Shukla, R.M., Pranggono, B., 2024. Transparency and privacy: the role of explainable AI and federated learning in financial fraud detection. *IEEe Access.* 12, 64551–64560. <https://doi.org/10.1109/ACCESS.2024.3394528>.
- Balsiger, D., Dimmler, H.-R., Egger-Horstmann, S., Hanne, T., 2024. Assessing large language models used for extracting table information from Annual Financial Reports. *Computers* 13. <https://doi.org/10.3390/computers13100257>.
- Cao, Y., Zhai, J., 2023. Bridging the gap—the impact of ChatGPT on financial research. *J. Chin. Econ. Bus. Stud.* 21, 177–191. <https://doi.org/10.1080/14765284.2023.2212434>.
- Çelik, T.B., İcan, Ö., Bulut, E., 2023. Extending machine learning prediction capabilities by explainable AI in financial time series prediction[Formula presented]. *Appl. Soft. Comput.* 132. <https://doi.org/10.1016/j.asoc.2022.109876>.
- Chen, X.-Q., Ma, C.-Q., Ren, Y.-S., Lei, Y.-T., Huynh, N.Q.A., Narayan, S., 2023. Explainable artificial intelligence in finance: a bibliometric review. *Financ. Res. Lett.* 56, 104145. <https://doi.org/10.1016/j.frl.2023.104145>.
- Cheng, Y., Zeng, Y., Zou, J., 2024. Harnessing ChatGPT for predictive financial factor generation: a new frontier in financial analysis and forecasting. *Br. Account. Rev.* <https://doi.org/10.1016/j.bar.2024.101507>.
- Chiu, I.-C., Hung, M.-W., 2025. Finance-specific large language models: advancing sentiment analysis and return prediction with LLaMA 2. *Pac. Basin Financ. J.* 90. <https://doi.org/10.1016/j.pacfin.2024.102632>.
- de Lange, P.E., Melsom, B., Vennerød, C.B., Westgaard, S., 2022. Explainable AI for credit assessment in banks. *J. Risk. Financ. Manag.* 15, 556. <https://doi.org/10.3390/jrfm15120556>.
- Delgadillo, J., Kinyua, J., Mutigwe, C., 2024. FinSoSent: advancing financial market sentiment analysis through pretrained large language models. *Big. Data Cogn. Comput.* 8. <https://doi.org/10.3390/bdcc8080087>.
- Dogra, V., Alharithi, F.S., Alvarez, R.M., Singh, A., Qahtani, A.M., 2022. NLP-based application for analyzing private and public banks stocks reaction to news events in the Indian Stock Exchange. *Systems (Basel)* 10. <https://doi.org/10.3390/systems10060233>.
- Dong, M.M., Stratopoulos, T.C., Wang, V.X., 2024. A scoping review of ChatGPT research in accounting and finance. *Int. J. Account. Inf. Syst.* 55, 100715. <https://doi.org/10.1016/j.accinf.2024.100715>.
- Dowling, M., Lucey, B., 2023. ChatGPT for (Finance) research: the Bananarama conjecture. *Financ. Res. Lett.* 53. <https://doi.org/10.1016/j.frl.2023.103662>.
- Eulerich, M., Sanatizadeh, A., Vakilzadeh, H., Wood, D.A., 2024. Is it all hype? ChatGPT's performance and disruptive potential in the accounting and auditing industries. *Rev. Account. Stud.* 29, 2318–2349. <https://doi.org/10.1007/s11142-024-09833-9>.
- Faccia, A., McDonald, J., George, B., 2024. NLP sentiment analysis and accounting transparency: a new era of financial record keeping. *Computers* 13. <https://doi.org/10.3390/computers13010005>.
- Fairhurst, D., Greene, D., 2024. How much does ChatGPT know about finance? *Financ. Anal. J.* 81, 12–32. <https://doi.org/10.1080/0015198X.2024.2411941>.
- Fatouros, G., Metaxas, K., Soldatos, J., Kyriazis, D., 2024. Can large language models beat wall street? Evaluating GPT-4's impact on financial decision-making with MarketSenseAI. *Neural Comput. Appl.* <https://doi.org/10.1007/s00521-024-10613-4>.
- Feuerriegel, S., Hartmann, J., Janiesch, C., Zszech, P., 2024. Generative ai. *Bus. Inf. Syst. Eng.* 66, 111–126.
- Fieberg, C., Hornuf, L., Streich, D., 2023. Using GPT-4 for financial advice. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.4499485>.
- Fior, J., Cagliero, L., Garza, P., 2022. Leveraging explainable AI to support cryptocurrency investors. *Futur. Internet* 14. <https://doi.org/10.3390/fi14090251>.
- Gao, W., Ju, M., 2024. A preliminary test of ChatGPT's ESG literacy. *Manag. Financ.* <https://doi.org/10.1108/MF-04-2024-0315>.
- Goodell, J.W., Kumar, S., Lim, W.M., Pattnaik, D., 2021. Artificial intelligence and machine learning in finance: identifying foundations, themes, and research clusters from bibliometric analysis. *J. Behav. Exp. Financ.* 32, 100577. <https://doi.org/10.1016/j.jbef.2021.100577>.
- Haugland Sundkvist, C., Kulset, E.M., 2024. Teaching accounting in the era of ChatGPT – the student perspective. *J. Account. Educ.* 69, 100932. <https://doi.org/10.1016/j.jaccedu.2024.100932>.
- Huang, A.H., Wang, H., Yang, Y., 2023. FinBERT: a large language model for extracting information from financial text. *Contemp. Account. Res.* 40, 806–841. <https://doi.org/10.1111/1911-3846.12832>.
- Hussain, S.M., Ahmad, N., Ahmed, S., 2023. Applications of high-frequency data in finance: a bibliometric literature review. *Int. Rev. Financ. Anal.* 89. <https://doi.org/10.1016/j.irfa.2023.102790>.
- Jha, M., Qian, J., Weber, M., Yang, B., 2024. ChatGPT and corporate policies. *Natl. Bureau Econ. Res.*
- Jiang, J., Kelly, B.T., Xiu, D., 2022. *Expected returns and large language models. Available SSRN.*
- Khan, F., Bartáková, G.P., Almador, A., Qayyum, A., Abeer, K., Durrani, A., 2025. Evaluating the capacity and limitations of generative AI in financial decision making. *Comput. Stand. Interfaces.* 93. <https://doi.org/10.1016/j.csi.2024.103965>.
- Khan, M.S., Umer, H., 2024. ChatGPT in finance: applications, challenges, and solutions. *Heliyon* 10. <https://doi.org/10.1016/j.heliyon.2024.e24890>.
- Ko, H., Lee, J., 2024. Can ChatGPT improve investment decisions? From a portfolio management perspective. *Financ. Res. Lett.* 64. <https://doi.org/10.1016/j.frl.2024.105433>.
- Lan, Y., Wu, Y., Xu, W., Feng, W., Zhang, Y., 2024. Chinese fine-grained financial sentiment analysis with large language models. *Neural Comput. Appl.* <https://doi.org/10.1007/s00521-024-10603-6>.
- Lee, D.K.C., Guan, C., Yu, Y., Ding, Q., 2024. A comprehensive review of generative AI in finance. *FinTech* 3, 460–478. <https://doi.org/10.3390/fintech3030025>.
- Leippold, M., 2023. Sentiment spin: attacking financial sentiment with GPT-3. *Financ. Res. Lett.* 55. <https://doi.org/10.1016/j.frl.2023.103957>.
- Li, H., Vasarhelyi, M.A., 2024. Applying large language models in accounting: a comparative analysis of different methodologies and off-the-shelf examples. *J. Emerg. Technol. Account.* 21, 133–152. <https://doi.org/10.2308/JETA-2023-065>.
- Li, X., Feng, H., Yang, H., Huang, J., 2024. Can ChatGPT reduce human financial analysts' optimistic biases? *Econ. Polit. Stud.* 12, 20–33. <https://doi.org/10.1080/20954816.2023.2276965>.
- Liu, L.X., Sun, Z., Xu, K., Chen, C., 2024. AI-driven financial analysis: exploring ChatGPT's capabilities and challenges. *Int. J. Financ. Stud.* 12. <https://doi.org/10.3390/ijfs12030060>.
- Mathebula, M., Modupe, A., Marivate, V., 2024. ChatGPT as a text annotation tool to evaluate sentiment analysis on South African financial institutions. *IEEe Access.* 12, 144017–144043. <https://doi.org/10.1109/ACCESS.2024.3464374>.
- Meng, K., Mahapatra, M.S., Xiao, J.J., 2025. Artificial Intelligence and consumer financial behavior: a systematic literature review and agenda for future research. *J. Consum. Behav.* 24, 1755–1786. <https://doi.org/10.1002/cb.2497>.
- Migliorini, S., 2024. China's Interim measures on generative AI: origin, content and significance. *Comput. Law Secur. Rev.* 53, 105985. <https://doi.org/10.1016/J.CLSR.2024.105985>.
- Moharrak, M., Mogaji, E., 2024. Generative AI in banking: empirical insights on integration, challenges and opportunities in a regulated industry. *Int. J. Bank Mark.* <https://doi.org/10.1108/IJBM-08-2024-0490>.
- Muneeb, D., Nobanee, H., Kamal, M.M., Shanti, H.Z., 2025. A bibliometric review of supply chain finance and digitalisation: mapping, current streams, and future research agenda. *Manag. Rev. Q.* 75, 43–81. <https://doi.org/10.1007/s11301-023-00374-0>.
- Muzanenhamo, P., Power, S.B., 2024. ChatGPT and accounting in African contexts: amplifying epistemic injustice. *Crit. Perspect. Account* 99. <https://doi.org/10.1016/j.cpa.2024.102735>.
- Niszczota, P., Abbas, S., 2023. GPT has become financially literate: insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice. *Financ. Res. Lett.* 58. <https://doi.org/10.1016/j.frl.2023.104333>.
- Nobel, S.M.N., Sultana, S., Singha, S.P., Chaki, S., Mahi, M.J.N., Jan, T., Barros, A., Whaiduzzaman, M., 2024. Unmasking banking fraud: unleashing the power of machine learning and explainable AI (XAI) on imbalanced data. *Inf* 15. <https://doi.org/10.3390/info15060298>.

- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Laltu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. <https://doi.org/10.1136/bmj.n71>.
- Pandey, D.K., 2025. FinTech literature reviews: a hybrid approach. *Financ. Res. Lett.* 79, 107249. <https://doi.org/10.1016/j.frl.2025.107249>.
- Pandey, D.K., Hassan, M.K., Kumari, V., Zaid, Y.Ben, Rai, V.K., 2024. Mapping the landscape of FinTech in banking and finance: a bibliometric review. *Res. Int. Bus. Financ.* 67. <https://doi.org/10.1016/j.ribaf.2023.102116>.
- Pelster, M., Val, J., 2024. Can ChatGPT assist in picking stocks? *Financ. Res. Lett.* 59. <https://doi.org/10.1016/j.frl.2023.104786>.
- Pfeifer, M., Marohl, V.P., 2023. CentralBankRoBERTa: a fine-tuned large language model for central bank communications. *J. Financ. Data Sci* 9. <https://doi.org/10.1016/j.jfds.2023.100114>.
- Quintais, J.P., 2025. Generative AI, copyright and the AI Act. *Comput. Law Secur. Rev.* 56, 106107. <https://doi.org/10.1016/J.CLSR.2025.106107>.
- Roberts, M.E., Stewart, B.M., Tingley, D., 2019. Stm: an R package for structural topic models. *J. Stat. Softw.* 91. <https://doi.org/10.18637/jss.v091.i02>.
- Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., Rand, D.G., 2014. Structural topic models for open-ended survey responses. *Am. J. Pol. Sci.* 58, 1064–1082. <https://doi.org/10.1111/ajps.12103>.
- Rodriguez, M.Y., Storer, H., 2020. A computational social science perspective on qualitative data exploration: using topic models for the descriptive analysis of social media data\*. *J. Technol. Hum. Serv.* 38, 54–86. <https://doi.org/10.1080/15228835.2019.1616350>.
- Roumeliotis, K.I., Tselikas, N.D., Nasiopoulos, D.K., 2024. LLMs and NLP models in cryptocurrency sentiment analysis: a comparative classification study. *Big. Data Cogn. Comput.* 8. <https://doi.org/10.3390/bdcc8060063>.
- Saxena, A., Santhanavijayan, A., Shakya, H.K., Kumar, G., Balusamy, B., Benedetto, F., 2024. Nested sentiment analysis for ESG impact: leveraging FinBERT to predict market dynamics based on eco-friendly and non-eco-friendly product perceptions with explainable AI. *Mathematics* 12. <https://doi.org/10.3390/math12213332>.
- Schimanski, T., Reding, A., Reding, N., Bingler, J., Kraus, M., Leipold, M., 2024. Bridging the gap in ESG measurement: using NLP to quantify environmental, social, and governance communication. *Financ. Res. Lett.* 61. <https://doi.org/10.1016/j.frl.2024.104979>.
- Schlosky, M.T.T., Karadas, S., Raskie, S., 2024. ChatGPT, help! I Am in financial trouble. *J. Risk. Financ. Manage.* 17. <https://doi.org/10.3390/jrfm17060241>.
- Shin, M., Ok, M.W., Choo, S., Hossain, G., Bryant, D.P., Kang, E., 2023. A content analysis of research on technology use for teaching mathematics to students with disabilities: word networks and topic modeling. *Int. J. Stem Educ.* 10. <https://doi.org/10.1186/s40594-023-00414-x>.
- Thanathamthee, P., Sawangarereak, S., Chantamunee, S., Nizam, D.N.M., 2024. SHAP-instance weighted and anchor explainable AI: enhancing XGBoost for financial fraud detection. *Emerg. Sci. J.* 8, 2404–2430. <https://doi.org/10.28991/ESJ-2024-08-06-016>.
- Trozee, A., Davies, T., Kleinberg, B., 2024. Large language models in cryptocurrency securities cases: can a GPT model meaningfully assist lawyers? *Artif. Intell. Law.* <https://doi.org/10.1007/s10506-024-09399-6>.
- Ullah, R., Ismail, H.Bin, Islam Khan, M.T., Zeb, A., 2024. Nexus between chat GPT usage dimensions and investment decisions making in Pakistan: moderating role of financial literacy. *Technol. Soc.* 76. <https://doi.org/10.1016/j.techsoc.2024.102454>.
- Wang, H., 2024. Decoding herding dynamics in the generative AI investment amid key technological advancements: a timeline perspective. *Financ. Res. Lett.* 64. <https://doi.org/10.1016/j.frl.2024.105432>.
- Wang, Q., 2025. Generative AI-assisted evaluation of ESG practices and information delays in ESG ratings. *Financ. Res. Lett.* 74. <https://doi.org/10.1016/j.frl.2025.106757>.
- Wood, D.A., Achhpilia, M.P., Adams, M.T., Aghazadeh, S., Akinyele, K., Akpan, M., Allee, K.D., Allen, A.M., Almer, E.D., Ames, D., Arity, V., Barr-Pulliam, D., Basoglu, K.A., Belnap, A., Bentley, J.W., Berg, T., Berglund, N.R., Berry, E., Bhandari, A., Bhuyan, M.N.H., Black, P.W., Blondeel, E., Bond, D., Bonrath, A., Borthick, A.F., Boyle, E.S., Bradford, M., Brandon, D.M., Brazel, J.F., Brockbank, B.G., Burger, M., Byzalov, D., Cannon, J.N., Caro, C., Carr, A.H., Cathey, J., Cating, R., Charron, K., Chavez, S., Chen, J., Chen, J.C., Chen, J.W., Cheng, C., Cheng, X., Christensen, B.E., Church, K.S., Cicone, N.J., Constance, P., Cooper, L.A., Correia, C.L., Coyne, J., Cram, W.A., Curtis, A., Daigle, R.J., Dannemiller, S., Davenport, S.A., Dawson, G.S., De Meyst, K.J.L., Dell, S., Demirkan, S., Denison, C.A., Desai, H., DeSimone, S., Diehl, L.M., Dimes, R., Dong, B., Donnelly, A., du Pon, A., Duan, H.K., Duffey, A., Dunn, R.T., Durkin, M.P., Dzurinin, A.C., Eberle, R.M., Ege, M.S., El Mahdy, D., Esplin, A., Eulerich, M., Everaert, P., Farah, N., Farish, L., Favere-Marchesi, M., Fayard, D., Filosa, J.R., Ford, M., Franz, D.R., Fulmer, B.P., Fulmer, S., Furner, Z.Z., Gantman, S., Garner, S., Garrett, J., Geng, X., Golden, J., Goldman, W., Gomez, J., Gooley, M., Granitto, S.P., Green, K.Y., Greenman, C.L., Gupta, G., Guymon, R.N., Hale, K., Harper, C.J., Hartt, S.A., Hawk, H., Hawkins, S.R., Hawkins, E.M., Hay, D.C., Heinzelmann, R., Henderson, C. D., Hendricks, B.E., Heninger, W.G., Hill, M.S., Holden, N., Holderness, D.K., Holt, T.P., Hoopes, J.L., Hsieh, S.-F., Huang, H.-W., Huang, T.-C., Huels, B. W., Hunter, K., Hurley, P.J., Inger, K., Islam, S., Ison, I., Issa, H., Jackson, A.B., Jackson, S.C., Janvrin, D.J., Jimenez, P.D., Johanson, D., Judd, J.S., Kawada, B.S., Kelton, A.S., Kern, S., Kerr, J.N., Keune, M.B., Kim, M., Knox, B.D., Kogan, G., Kotb, A., Krane, R., Kremin, J., Krieg, K.S., Kugel, J., Kulset, E.M., Kuruppu, C., LaDuca, G., Lambertson, B.A., Lamboy-Ruiz, M.A., Lang, B., Larocque, S.A., Larson, M.P., Lawson, B.P., Lawson, J.G., Lee, L., Lenk, M.M., Li-Kuehne, M., Liljegren, J., Lin, Y.-H., Liu, W.-P., Liu, Z., Lock, B., Long, J.H., Loraas, T., Lowensohn, S., Loy, T.R., Lyngstadaas, H., Maas, W., MacGregor, J.E., Madsen, D.O., Malone, C.L., Margolin, M., Marshall, M.E., Martin, R.M., Mpofu, C.M., McCoy, C., McGuigan, N.C., McSwain, D.N., Meckfessel, M.D., Mellon, M.J., Melton, O.S., Mercado, J.M., Mitsuda, S., Modugu, K., Moehle, S., Chaghervand, A.M., Moffitt, K., Moon, J.S., Muehlmann, B., Murray, J., Mwaungulu, E.S., Myers, N., Naegle, J.C., Ndicu, M.J., Nelson, A.S., Nguyen, A.L., Niederkofler, T., Nikbakht, E., O'Brien, A.D., Ogunade, K.M., O'Leary, D., Oler, M.J., Oler, D.K., Olsen, K.J., Otolal, J.I., Outlaw, K.W., Ozlanski, M.E., Parlier, J., Paterson, J.S., Pearson, C.A., Petersen, M.J., Petra, S.T., Pickard, M.D., Pickerd, J., Pinsker, R., Plante, C., Plečnik, J.M., Price, R.A., Quick, L.A., Raedy, J., Raschke, R., Ravenscraft, J., Richardson, V., Rixom, B.A., Robertson, J.F., Rock, I., Romney, M.A., Rozario, A., Ruff, M.F., Rupley, K., Saedi, A., Saiewitz, A., Salzsieder, L.W., Sarkar, S., Sauls, M., Scanlan, T.A., Schaefer, T.J., Schaupt, D., Schneider, G.P., Seebeck, A., Sellers, R.D., Seto, S.C., Sevel, R.-L., Shan, Y., Sherwood, M.G., Singorahardjo, M., Skaftadottir, H.K., Skomra, J., Smith, J.L., Smith, D.O., Smith, J., Snow, M.C., Sommerfeldt, R.D., Sorensen, K.B., Sorensen, T.L., Spieler, A.C., Stallings, M.A., Stallings, L., Stancill, A., Stanley, J.D., Stefaniak, C.M., Stephens, N.M., Stewart, B. W., Stratopoulos, T.C., Street, D.A., Subedi, M., Summers, S.L., Sundkvist, C.H., Synn, C., Tadesse, A., Tapis, G.P., Tassin, K., Taylor, S., Teal, M., Teeter, R., Tharapos, M., Theis, J.C., Thomas, J., Thompson, K.S., Thornock, T.A., Tietz, W., Trivalent, A.M., Trinkle, B.S., Truelson, J.M., Turner, M.C., Vagner, B., Vakilzadeh, H., van der Geest, J., van Pelt, V., Vandervelde, S.D., Vega, J., Vera-Muñoz, S., Villanueva, B., Vincent, N.E., Wagener, M., Walton, S., Warne, R.C., Watanabe, O.V., Watson, D., Watson, M.W., Weber, J., Weirich, T., West, A.N., Wilford, A.L., Wilson, A.B., Winrow, B., Winrow, T., Winrow, T.S., Wiseman, D., Witte, A.L., Wood, B.D., Wood, J., Woolley, D., Wright, N.S., Wu, J., Xiong, X., Yatsenko, D., Yazzie, C.E., Young, G.M., Zhang, C., Zimmerman, A.B., Zoet, E., 2023. The ChatGPT artificial intelligence chatbot: how well does it answer accounting assessment questions? *Issues account. Educ* 38, 81–108. <https://doi.org/10.2308/ISSUES-2023-013>.
- Wood, K., Pyun, C., Pham, H., 2025. Beyond green labels: assessing Mutual funds' ESG commitments through large language models. *Financ. Res. Lett.* 74. <https://doi.org/10.1016/j.frl.2024.106713>.
- Zhao, J., Wang, X., 2024. Unleashing efficiency and insights: exploring the potential applications and challenges of ChatGPT in accounting. *J. Corp. Account. Financ.* 35, 269–276. <https://doi.org/10.1002/jcaf.22663>.
- Zheng, X., Li, J., Lu, M., Wang, F.-Y., 2024. New paradigm for economic and financial research with generative AI: impact and perspective. *IEEe Trans. Comput. Soc. Syst.* 11, 3457–3467. <https://doi.org/10.1109/TCSS.2023.3334306>.