المجلة الدولية للأداء الاقتصادي
International journal of economic performance
ISSN: 2661-7161 EISSN:2716-9073
Publications are Open Access in this journal
journal home page : https://ijep.dz/index.php/IJEP/index

# Large Language Models in Digital Banking: A Systematic Review and Implications for Financial Inclusion (2015–2025)

Mohamed Djafar Henni *
mhenni@iu.edu.sa, Department of Economics, College of Business, Islamic University of Madinah, Saudi Arabia

**ARTICLE INFO**

**Article history**:
Received : 21/04/2026
Accepted : 30/05/2026
Published:30/06/2026

**Keywords**: Large Language Models; Banking Management; Generative AI; Fintech; Systematic Review; Risk Management and Financial Prediction

**Jel Codes:**
M20; P23; F10; H23

**ABSTRACT**

The rapid emergence of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) is fundamentally transforming the financial sector by enabling advanced capabilities in processing large volumes of unstructured textual data. While previous studies have explored broad applications of artificial intelligence in finance, a dedicated synthesis focusing specifically on banking management and its implications for digital banking transformation remains limited. In response, this article provides a systematic state-of-the-art review of 91 studies published between 2015 and 2025, identified through the Web of Science and Scopus databases using a PRISMA-based methodology.The review synthesizes the literature across five key domains: (1) Policy Interpretation and Sentiment Analysis, (2) Risk Management and Financial Prediction, (3) Regulatory Technology (RegTech) and Compliance, (4) Operational Efficiency and Process Management, and (5) Customer-Facing Applications and Advisory Services. Bibliometric evidence reveals a rapid acceleration of scholarly interest beginning in 2023, with Risk Management and Financial Prediction representing the most prominent research stream (28.7% of publications). The findings demonstrate how LLM-driven tools are redefining traditional banking management practices by enabling contextual interpretation of central bank communications, improving financial risk forecasting, automating regulatory compliance processes, and enhancing operational decision-making. digital banking services in improving customer interaction, expanding access to financial information, and potentially supporting broader financial inclusion within increasingly digitalized financial systems. Finally, the study identifies critical limitations in the existing literature particularly regarding data privacy, model interpretability, and geographic bias and proposes a strategic roadmap outlining future research directions for responsible and inclusive AI adoption in banking.



* *Corresponding author, E-mail address:* mhenni@iu.edu.sa *Tel :0534087690*

## 1. Introduction

The accelerated and potent advent of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) is drastically reshaping the financial domain. This is because GenAI and LLMs hold unparalleled potential for handling and understanding vast amounts of data or information, thus causing a complete overhaul of the very foundation of financial stability, risk management, and the markets themselves (Aldasoro et al., 2025; Coelho E Silva et al., 2024; Eisfeldt & Schubert, 2025). The financial sector, where vast streams of information are continuously synthesized into market price signals, has shown strong adoption of LLMs due to their superior contextual reasoning and high-quality content generation capabilities (Haidar & Abbass, 2025). This technology has emerged as a disruptive force across diverse domains, including corporate finance, asset pricing, household finance, banking management, and regulatory technology (Feng et al., 2025; Mo & Ouyang, 2025). As this landscape continues to evolve, the effectiveness of banking management has become increasingly vital for maintaining global financial stability (Cao & Feinstein, 2024; Fan, 2024; Sideras et al., 2024). Simultaneously, the integration of sophisticated computational tools is redefining core banking processes, extending from transactional activities and customer support to credit assessment and interpreting Regulatory Technology (RegTech) (Saxena et al., 2024).

However, aside from the changes brought about by the transformation of institutions and sectors, the spread of GenAI and LLMs also has macroeconomic implications. In fact, recent systemic studies, such as "Integrating Digital and AI-Driven Productivity into National Accounts: A Systemic Analysis of Economic Impacts in Emerging and Advanced Economies," suggest that the productivity brought about by AI disrupts the traditional framework of national accounting, especially with regard to digital capital, intangible assets, and algorithmic decision-making systems. As the banking industry, which is highly dependent on AI, becomes more and more reliant on data-driven automation and intelligent systems, the traditional measure of productivity may fall short in estimating the actual economic value created. Therefore, the adoption of AI in banking management should not only be understood within its operational and managerial frameworks but also within the context of the overall shift in aggregate productivity and structural economic change in emerging and advanced economies(Mohamed, Henni, and Sorour 2026).

A significant portion of recent research has focused on applying LLMs and Natural Language Processing (NLP) techniques to analyze financial textual data. This trend is particularly visible in studies of central bank communication and monetary policy, where researchers applied models such as FinBERT and specialized LLMs (e.g., CentralBankRoBERTa) to measure sentiment, predict policy uncertainty, classify emotions, and assess the consistency of official communications (Ardekani et al., 2024; Bertsch et al., 2025; Cho & Jung, 2026; Ito et al., 2025; Kanelis & Siklos, 2025; Patterson, 2025; Pfeifer & Marohl, 2023). Further applied research has examined the transformative role of GenAI in specific banking areas, such as enhancing credit default prediction by refining loan assessment texts (Wu et al., 2025), evaluating investor risk profiling capabilities (Hens & Nordlie, 2025), detecting corruption through textual analysis

(Damiano et al., 2025), and identifying customer intent in data-limited banking domains (Loukas et al., 2023; Srivastava, 2024). Research also explored the organizational implications of AI, such as negative stock market reaction of US banks following ChatGPT's launch (Beckmann & Hark, 2024) and how AI-generated insights can improve financial management in enterprises with limited resources (Metzger et al., 2025). These studies clearly demonstrate that banking is leading the way in AI adoption in finance.

Although AI applications in the financial and banking sectors are widely studied, no comprehensive review dedicated specifically to banking management has yet emerged. To the author's knowledge, existing reviews either discuss broad applications of GenAI or LLMs in finance (Haidar & Abbass, 2025; Lee et al., 2024; Saxena et al., 2024), or concentrate on specific financial markets or macroeconomic issues (Iadisernia & Camassa, 2025; Mahendran et al., 2025). As a result, scholars and practitioners lack a targeted synthesis on management, strategy, operations, and risk considerations unique to banking in the AI era.

To fill this critical gap, the present review article undertakes a rigorous, targeted analysis of research focusing on banking management in finance. This study offers a uniquely timely perspective by analyzing both recently published and *forthcoming* articles (i.e., scheduled for publication in 2026 across leading journals). By synthesizing these foundational and emerging works, the review aims to provide an unprecedented roadmap of the opportunities, challenges, and strategic approaches in modern banking management. Following the introduction (Section 1), Section 2 details the Search method procedure used to identify relevant papers across the Web of Science and Scopus databases. Section 3 offers a historical outline of LLMs and their evolution, thereby establishing the technological context for the present review. Section 4 provides an overview of the number of publications and presents a keyword co-occurrence analysis of the research landscape through quantitative trends. The core of the review, Section 5, synthesizes the findings across five primary application domains: Policy Interpretation and Sentiment, Risk Management and Financial Prediction, Regulatory Technology (RegTech) and Compliance, Operational Efficiency and Process Management, and Customer-Facing Applications and Advisory Services. Finally, Section 6 presents the Conclusion and Future Research Directions.

## 2. Search Method Procedure

In this section, we outline the methods that were used to find and acquire the academic journals needed for this survey. The methods used to find relevant academic journals for the survey are presented below.

### 2.1. Search Method

The primary platforms utilized for identifying relevant literature for this review are the Web of Science and Scopus databases. A comprehensive search strategy was employed using a list of keywords to capture the breadth of research at the intersection of LLMs and banking. The search involved combinations of terms related to language models (i.e., "Large Language

Models", "LLMs") paired with the keyword "bank". Table 1 presents several search queries used in this survey

Table 1. Search term combinations used in Web of Science and Scopus.

| | LLM / NLP Term | Finance Term | Example Query Fragment |
|---|---|---|---|
| **Web of Science** | "Large Language Models" | "bank*"[1] | TS=("Large Language Models" AND "bank*") |
| | "LLMs" | "bank*" | TS=("LLMs" AND "bank*") |
| **Scopus** | "Large Language Models" | "bank*" | TITLE-ABS-KEY=("Large Language Models" AND "bank*") |
| | "LLMs" | "bank*" | TITLE-ABS-KEY=("LLMs" AND "bank*") |

[1]: In Web of Science and Scopus, the asterisk (*) is used as a truncation operator to retrieve all word variants sharing the same root. For example, "bank*" captures bank, banks, banking, banker, etc.

With the aim of ensuring a wide and transparent scope of review of the literature selected for the systematic review, the selection of the literature for the systematic review was based on a systematic and phased approach of literature selection/filtering in compliance with the PRISMA guidelines (Moher et al., 2010). The process began with a broad, unrestricted query that retrieved over 1,000 studies related to LLMs and GenAI. Subsequently, the initial set of publications was filtered more towards focusing exclusively on English-language publications within the broad domains of economics and finance. Furthermore, the search was filtered towards publications within the range of 2015-2025, a period characterized by significant developments. This further filtered the search to yield a total of 53 publications from the Web of Science database and a additional 70 publications from the Scopus database, amounting to a total of 123 publications prior to the removal of duplicates. After the removal of duplicates, a final set of 91 publications was obtained, and they constituted the foundational set for the purposes of the review. The reference management software Mendeley (Elsevier, 2020) was used to organize the selected literature. Fig. 1 summarizes the PRISMA-based search and selection process employed in this study.

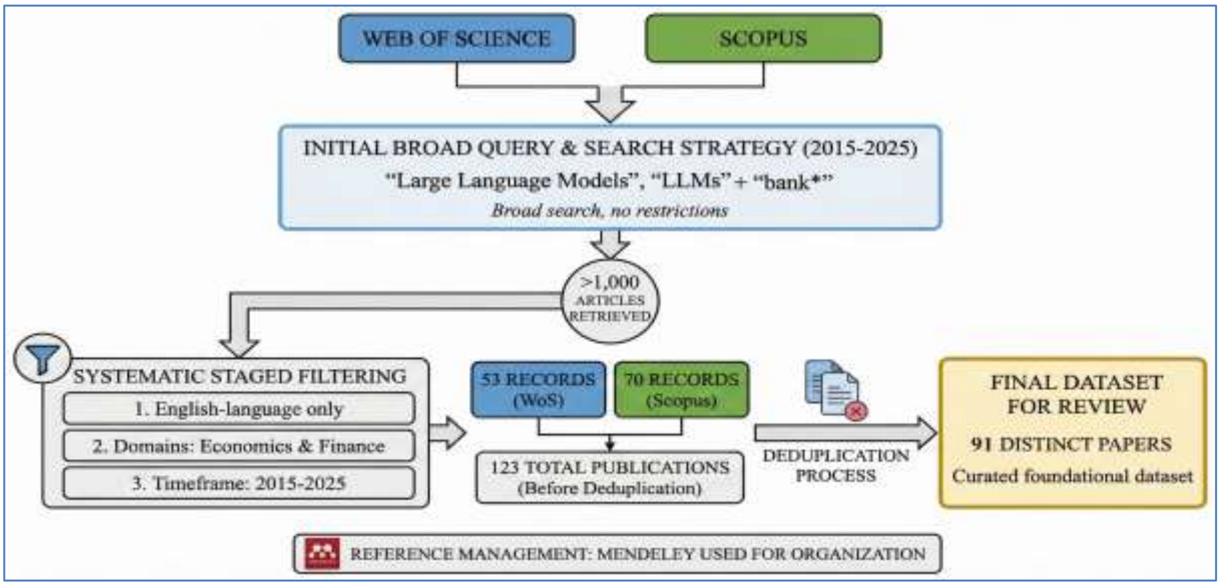


**Fig. 1 PRISMA-based search workflow to locate and retrieve academic articles and research paper**

## 2.2. Other Reviews

The landscape of financial technology research is rich with surveys detailing the applications of AI and LLMs; however, a notable gap remains in providing a dedicated, comprehensive review centered exclusively on banking management in the GenAI era. The available high-impact reviews cover financial domains, offering background information but have not resulted in focused synthesis appropriate as a guide for banking executives. For example, Bae et al. (2023) conducted a survey of the latest progress in methods and usage of textual analysis, specifically within accounting and finance literature and applications. It covered topics, among others, from corporate communications to investor relations, and offers guidelines on the selection of appropriate text analysis methods and best practices, and validation techniques and reporting of text-based evidence. This review served as a methodological foundation for utilizing unstructured financial text data.

Dong et al. (2024) carried out a critical analysis of the role of ChatGPT and LLMs in accounting and finance and found that the main issues regarding the applications of the aforementioned chatbot and their impact on accounting and finance can be described using the following three main issues: the application of the discussed technology in multiple issues of accounting and finance and their usage in the field of research as a tool for summarizing and classifying data. However, the critical issues have not been given due consideration.

The current debate on the influence of GenAI on finance is centered on six important issues identified in the research carried out by Ali et al. (2025) namely financial decision-making, ESG analytics, stock prediction, advanced modeling for detecting fraud and using explainable AI (xAI), the role of ChatGPT (OpenAI, 2024) in accounting and education, and the impact of sentiment analysis using domain-specific LLMs. The focus of the above issues highlights the positives of using GenAI in predicting and making financial decisions and in conducting ESG risk analysis. Additionally, the negative issues related to the application of GenAI are also discussed. Moreover, Dong et al. (2024)carried out a critical analysis of the role of ChatGPT and

LLMs in accounting and finance and found that the main issues regarding the applications of the aforementioned chatbot and their impact on accounting and finance can be described using the following three main issues: the application of the discussed technology in multiple issues of accounting and finance and their usage in the field of research as a tool for summarizing and classifying data. However, the critical issues have not been given due consideration. Eisfeldt and Schubert (2025) examined GenAI's disruptive role as a major technology shock in finance. The study revealed that financial occupations exhibit high exposure to the productivity effects of GenAI, and it offered practical advice and directions for future research.

Another review c productivity onducted by Kong et al. (2024) offered a broad overview of LLMs in finance and investment management by discussing recent models, technologies, opportunities, and challenges. The review summarized the power of LLMs in processing unstructured data for financial analysis while emphasizing critical challenges, such as issues with data quality, model complexities, and ethical concerns. Importantly, these challenges were framed not only as constraints but also as avenues for future research and innovation. Lee et al. (2024) presented a systematic review of the integration of GenAI into the financial sector. Through the use of advanced topic modeling, research was categorized into predominant themes.

A review of this nature was carried out by Mo and Ouyang in 2025. In this review, Mo and Ouyang (2025) conducted a comprehensive review of the intersection between AI/GenAI and financial economics. The authors synthesized the literature by organizing it into six core thematic areas: (i) the role of GenAI and LLMs as analytical tools, autonomous agents, and sources of external economic shocks; (ii) corporate finance; (iii) asset pricing; (iv) household finance; (v) labor economics; and (vi) risks associated with AI in financial markets. This structured synthesis provided a broad perspective on the economic implications of AI and its influence on financial market behavior. Saxena et al. (2024) specifically focused on the integration of GenAI within the broader Banking, Financial Services, and Insurance (BFSI) sector. It outlined GenAI applications across core banking processes and proposed a structured adoption roadmap. In addition, the authors explicitly acknowledged ethical and moral considerations associated with the deployment of GenAI in BFSI environments.

The future research, as collected from the Research Square platform, shows that forthcoming reviews will focus more on systematic, granular, and comprehensive studies, investigating in-depth information about particular LLM applications and fundamental methodologies within the financial industry. The systematic review carried out byUpcoming research retrieved from the Research Square database demonstrates a trend toward more systematic, granular, and comprehensive reviews, offering in-depth analyses of specific LLM applications and foundational methodologies within the financial sector. Raliphada et al. (2025) focused specifically on reviews of the application of Transformer-based NLP techniques in credit risk prediction. The authors investigated how techniques like BERT(Devlin, Chang, Lee, & Toutanova, 2018), RoBERTa (Liu et al., 2019), and LLaMA (Touvron et al., 2023) improve the predictive ability of financial models, as well as some significant shortcomings associated with their applications and usability. From the analysis of 63 studies, the authors have confirmed that models based on the transformer show tremendous potential in improving predictive accuracy in credit risk, especially when the temporal and sentiment features are incorporated.

Lengyel et al. (2025) conducted a systematic literature review to evaluate the fintech implications of integrating Machine Learning (ML) algorithms into cryptocurrency trading strategies. By analyzing 57 peer-reviewed studies and high-quality preprints published between 2016 and 2025, the authors aimed to address the fragmented understanding of how advanced computational models impact trading efficacy, risk management, and financial innovation. The findings revealed that deep reinforcement learning significantly improves predictive accuracy, profitability, and risk management.

Although the review papers provide extensive coverage of the technical capabilities and broad application domains of generative GenAI in the financial sector, a cohesive synthesis that explicitly addresses the strategic and operational management of banking institutions remains notably absent. Existing review studies tend to concentrate on three primary dimensions. First, *broad financial applications* are emphasized, including predicting stock market returns (Ali et al., 2025), volatility forecasting, and portfolio optimization (Kong et al., 2024). Second, *methodological advancements* receive substantial attention, particularly in areas such as synthetic data generation (Lee et al., 2024), mobilizing unstructured text (Bae et al., 2023), and designing agentic systems (Mo & Ouyang, 2025). Third, *narrowly scoped* applications are only briefly addressed, with regulatory compliance mentioned at a high level in the BFSI context (Saxena et al., 2024), and machine learning integration within cryptocurrency trading environments (Lengyel et al., 2025). *Fig. 2* summarizes the landscape of existing review articles compared to the current focus in banking management. Collectively, this existing and upcoming body of work confirms the central premise of the current review: no single existing or readily identifiable forthcoming review that offers the focused, systematic synthesis specifically addressing the strategic and operational management of banking institutions.

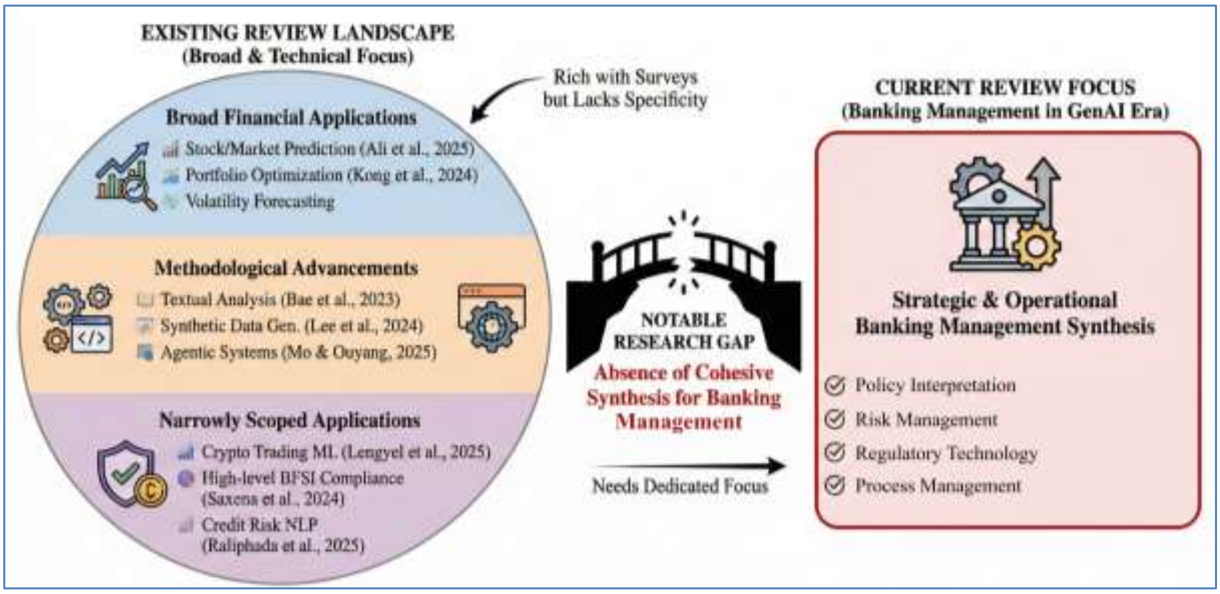


**Fig. 2** **Landscape of existing review articles vs. current banking management focus**

## 3. Large Language Models (LLMs)

The origins of language modeling lie in early computational linguistics and symbolic AI. In the 1950s, researchers such as Alan (1950 as cited in Muggleton, 2014) proposed theoretical frameworks for machines capable of simulating human conversation. Progress at the time, however, was limited by constrained computational resources and the scarcity of large-scale data. Early rule-based systems, exemplified by Joseph Weizenbaum's ELIZA (1966), relied on scripted responses and pattern matching rather than genuine language understanding. Later, the 1980s witnessed the emergence of statistical language models, which employed probabilistic methods to predict word sequences. N-gram models (i.e., models that estimate the likelihood of a word based on its preceding n-1 words) became a cornerstone of speech recognition and machine translation (Jelinek, 1976). These models, while effective for narrow tasks, lacked contextual depth and struggled with ambiguity.

The resurgence of neural networks in the 2000s marked a paradigm shift. Seminal work conducted by Bengio (Bengio et al., 2003) introduced neural language models capable of learning distributed representations of words, laying the foundation of word embeddings. Building on this concept, Mikolov et al. (2013) developed Word2Vec, a highly efficient embedding technique that maps words to vectors using shallow neural networks. Word2Vec's ability to analogize (e.g., "king – man + woman = queen") demonstrated the potential of neural methods for capturing linguistic patterns. Concurrently, the GloVe model (Pennington et al., 2014) provided a global context-aware embedding by combining the strengths of global matrix factorization and local context window methods. This model produced vector representations with meaningful substructure, hence enabling strong performance on word analogy, similarity, and named entity recognition tasks.

The mid-2010s marked the integration of deep learning into natural language processing (NLP). Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks allowed models to handle sequential data and retain contextual information. A sequence-to-sequence (Seq2Seq) architecture, introduced by Sutskever et al. (2014), paired encoder-decoder networks to transform input sequences into output sequences, thereby revolutionizing tasks like text summarization. Despite these advances, RNNs struggled to capture long-range dependencies due to the vanishing gradient problem. Attention mechanisms addressed this limitation by enabling models to selectively weight relevant parts of the input (Vaswani et al., 2017). This innovation led to the Transformer architecture, which replaced recurrence entirely with self-attention layers.

Several hybrid models bridged the gap between traditional and neural approaches. ELMo (Embeddings from Language Models), introduced by the Allen Institute for AI in 2018 (Peters et al., 2018), relied on bidirectional LSTMs to produce context-sensitive word embeddings. Its dynamic representations outperformed static embeddings such as Word2Vec and achieved state-of-the-art results in tasks, including question answering. In the same year, Transformer architecture sparked a revolution that enabled the development of modern LLMs. Two landmark models emerged: Google's BERT (Bidirectional Encoder Representations from Transformers) and OpenAI's GPT (Generative Pre-trained Transformer). GPT-1 (Radford & Narasimhan, 2018) demonstrated the effectiveness of unsupervised pre-training followed by task-specific fine-tuning. The model used a unidirectional Transformer decoder to predict the next word in a sequence after it was trained on the BookCorpus dataset (7,000 unpublished books). Despite its 117 million parameters, GPT-1 achieved strong performance on tasks like text classification and entailment. BERT (Devlin, Chang, Lee, Google, et al., 2018), on the other hand, introduced bidirectional context by training on masked language modeling (MLM) and Next Sentence Prediction (NSP). By predicting 15% of masked input tokens, BERT learned deep contextual relationships. This training strategy enabled BERT to outperform GPT-1 on the GLUE benchmark, achieving 80.2% relative to GPT-1's 72.8% (A. Wang et al., 2018).

The early success of Transformers triggered an accelerated pursuit of scale in both model parameters and training data. In 2019, OpenAI released GPT-2, a 1.5-billion-parameter model trained on 40GB of web text. GPT-2 demonstrated a strong capacity to generate coherent and contextually appropriate language, which raised concerns regarding potential misuse. Subsequently, GPT-3 marked a substantial advancement with 175 billion parameters, trained on 570GB of text from Common Crawl, books, and Wikipedia (Brown et al., 2020). Its few-shot learning capabilities enabled users to prompt the model with minimal examples, achieving human-like performance in areas such as writing, coding, and reasoning. The release of GPT-3 sparked widespread adoption of LLMs across numerous applications, including conversational agents and automated content generation systems. By 2020, LLMs had been integrated into various domains, including healthcare (e.g., assisting in medical documentation and literature reviews) (Benjamens et al., 2020), education (e.g., automating feedback and providing personalized tutoring) (Peng et al., 2020), Engineering (Y. Wang et al., 2020), and finance (Li et al., 2023).

## 4. Bibliometric Overview of LLMs in Banking Management

This section organizes and synthesizes the body of publications identified on the application of LLMs in banking management, following the search methodology outlined in Section 2. The temporal evolution of scholarly interest in LLMs within the banking management sector is illustrated in *Fig. 3*. As illustrated in *Fig. 3*, the chronological distribution highlights a dramatic shift in academic attention toward LLM-driven banking research. Specifically, the field has transitioned from occasional contributions, averaging fewer than five articles annually between 2015 and 2020, to a period of explosive growth. Although a modest increase is observed in 2021, with nine publications, a substantial inflection point emerged in 2023, when the number of studies rose to twelve, coinciding with the widespread adoption and mainstream recognition of GenAI technologies. This growth trajectory intensified markedly in subsequent years, with 32 publications in 2024 and 45 in 2025. This rapid escalation, visualized by the ascending bar chart and its intersecting trend line in *Fig. 3*, underscores the current transformative impact of LLM technologies on the banking sector. It also reflects a maturing research domain that has recently become the primary driver of contemporary bank management literature.

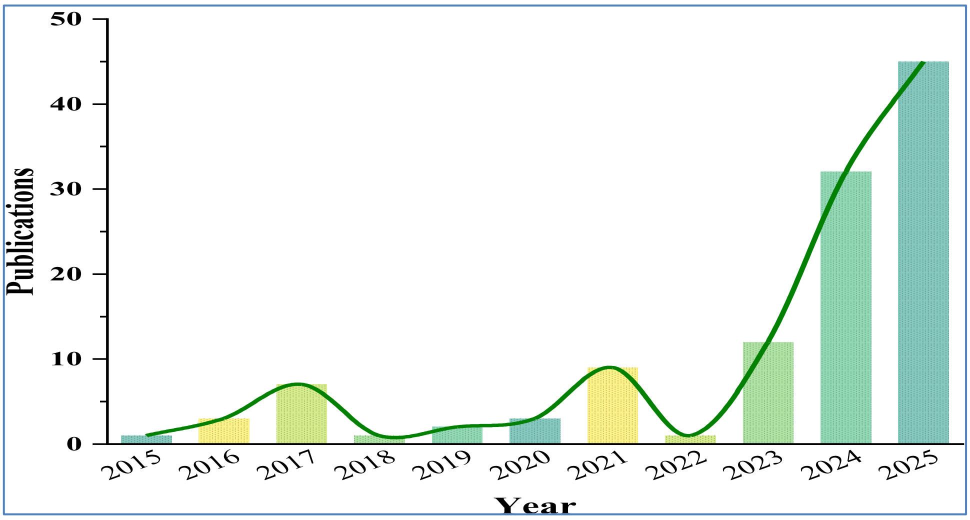


Fig. 3 **Total number of publications per year**

These reviewed articles are categorized based on their primary application area into five categories: Policy Interpretation and Sentiment, Risk Management and Financial Prediction, Regulatory Technology (RegTech) and Compliance, Operational Efficiency and Process Management, and Customer-Facing Applications and Advisory Services. The distribution of publications across these categories is presented in *Fig. 4*. As depicted in *Fig. 4*, Risk Management and Financial Prediction emerge as the most prominent themes, accounting for 28.7% of the articles. It highlights the industry's emphasis on stability and predictive accuracy amid heightened market volatility. This is followed closely by Operational Efficiency and Process Management (23.8%), reflecting sustained academic and industry interest in automating and optimizing internal banking operations. Collectively, Policy Interpretation and Sentiment (21.3%) and Regulatory Technology (RegTech) and Compliance (17.5%) accounted for more than a third of the research, highlighting the critical role of AI in navigating complex legal landscapes. Notably, despite the widespread association of digital banking transformation with enhanced customer experience, Customer-Facing Applications and Advisory Services represent the smallest share of the literature (8.8%).

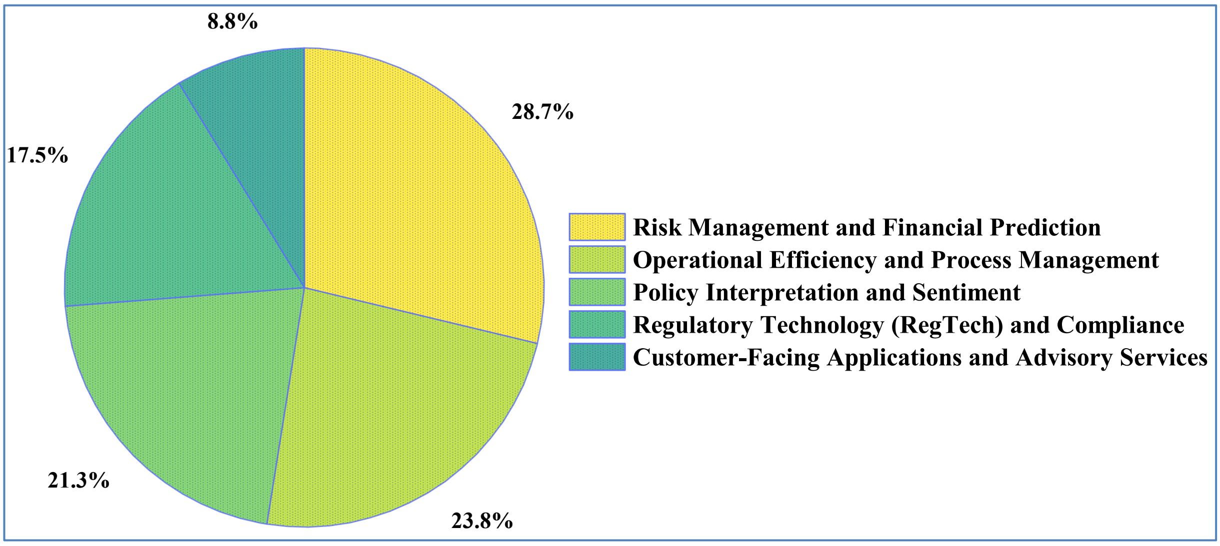


Fig. 4 **Percentage of publications in each subfield based on keyword searches**

*Fig. 5* illustrates the distribution of the top leading publishers represented in the reviewed literature. There is a strong concentration of research output within a small number of dominant academic publishers. Elsevier emerges as the leading publisher, accounting for 28 articles, followed by IEEE (14), Springer Nature (13), and ACM (12). Collectively, these four organizations publish over two-thirds of the total literature, reflecting their dominance in technical and financial research. The remaining articles are distributed among Emerald Group Publishing (10), Cambridge University Press (8), Wiley (6), Taylor & Francis (6), and MDPI (3). This distribution demonstrates a scholarly landscape that is both diverse and anchored by established high-impact publishers.

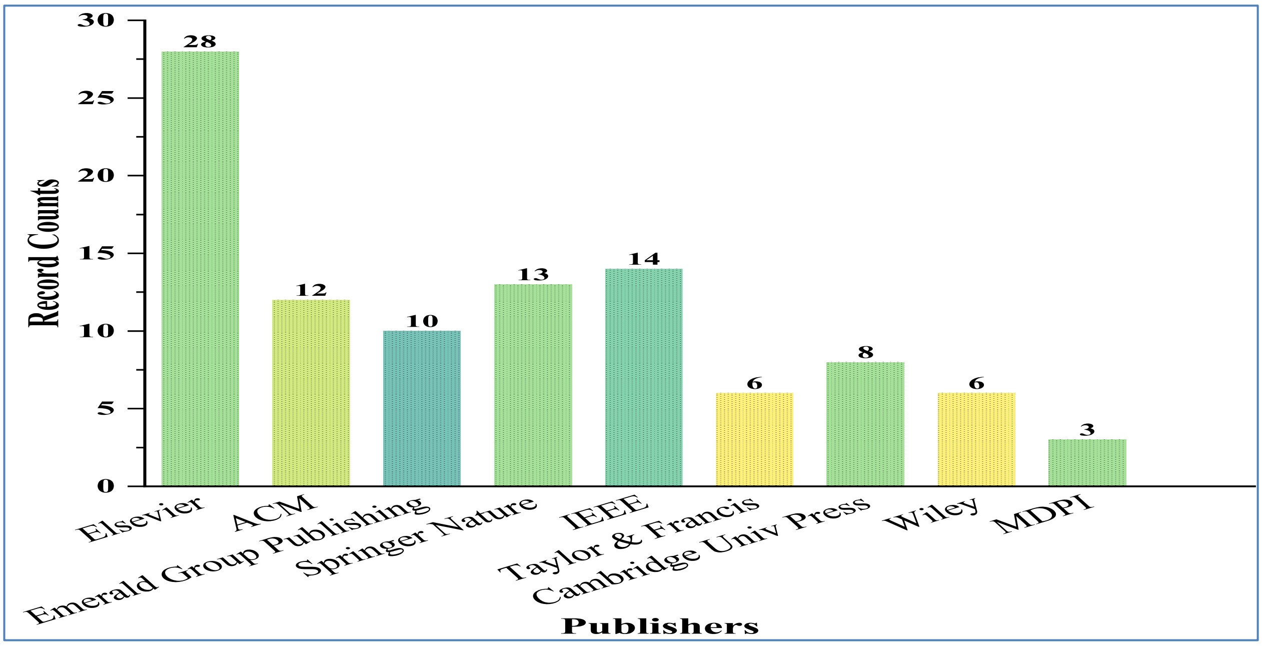


Fig. 5 **Number of publications across 10 leading publishers**

Moreover, the geographical distribution of research output by country is presented in *Fig. 6*, which displays only the top ten contributing nations based on the number of published articles. As shown, the United States leads the field with 32 publications, substantially surpassing other countries and serving as a central hub for AI-banking research. Europe represents a major geographical cluster, with the United Kingdom (15 articles) leading the region, followed by strong contributions from Germany (10), France (10), Spain (8), Italy (7), and Norway (6). This concentration suggests a well-established research infrastructure and significant regulatory engagement in European banking sectors. Furthermore, significant activity is observed in Asia, particularly from India (13) and China (8), which reflects the rapid digital transformation of financial services in these emerging markets. Canada completes the top contributors with 7 articles. Although only the most productive countries are displayed in *Fig. 6*, it should be noted that several other countries also contribute to the literature, reflecting the global and distributed nature of research activity in this domain.

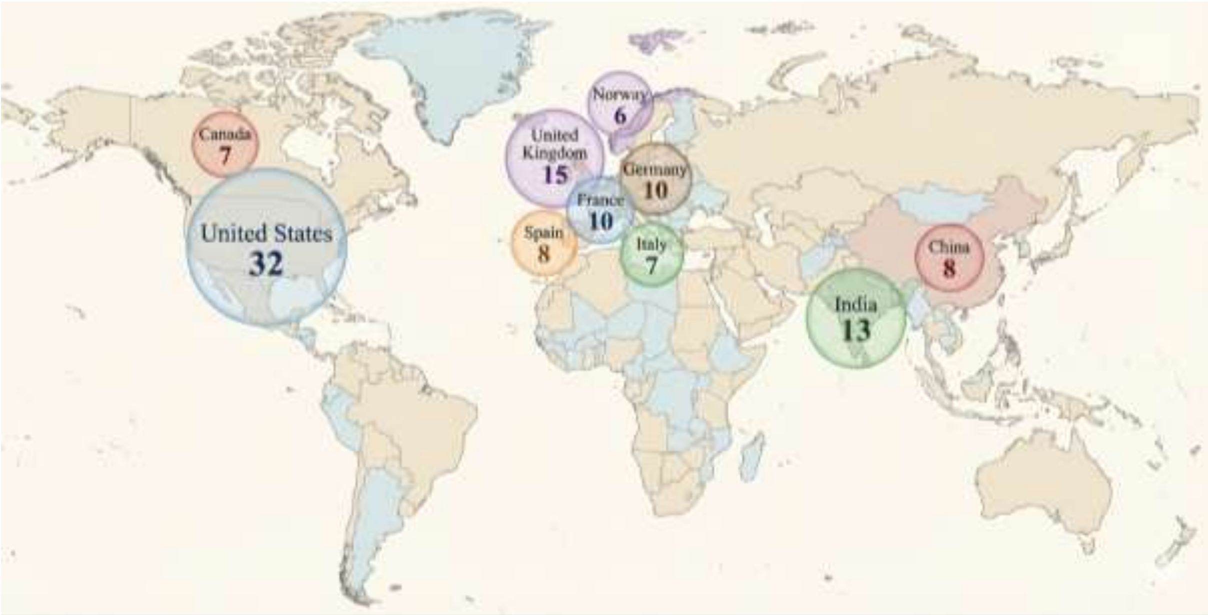


Fig. 6 **Geographic distribution of publications among the ten leading countries**

Fig. 7 depicts the collaboration network of 281 authors using a cluster-based visualization. At the macro level, individual authors are aggregated into clusters, with each cluster representing a distinct group of co-authors. The clustering effect indicates that co-authorship is distributed across smaller, localized groups rather than forming a globally interconnected network. A zoomed-in view complements this representation by displaying the individual nodes within each cluster, where node size corresponds to the number of publications contributed by each author. Further, the color-coded overlay indicates that the most significant collaborative efforts have occurred between 2022 and 2025.

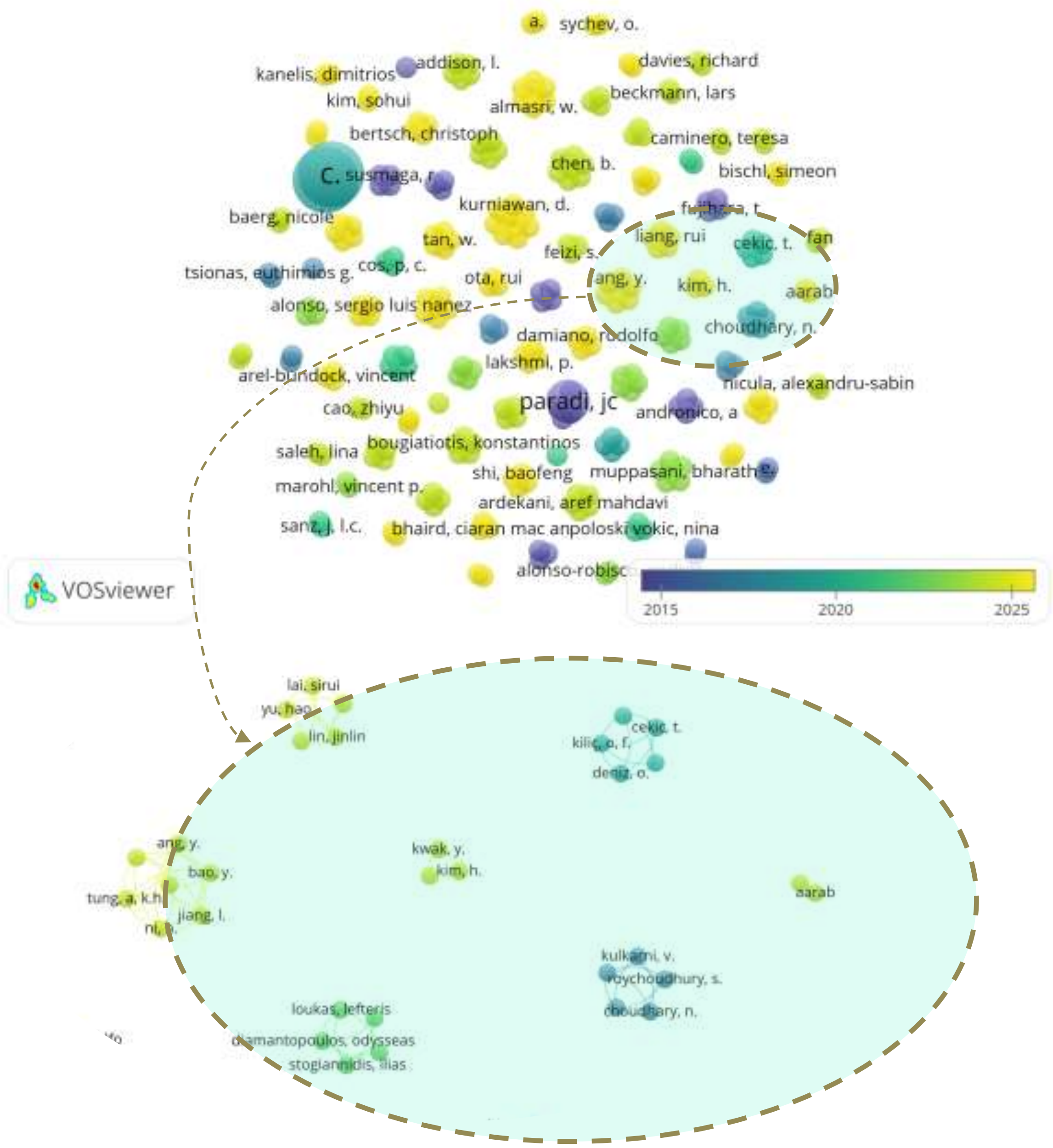


Fig. 7 Network map of co-author interactions within the academic community

A comprehensive network visualization of keyword co-occurrence in banking management, generated using VOSviewer (van Eck & Waltman, 2009), is illustrated in Fig. 8. In this network, each node on the network represents a specific keyword, while the connecting lines between the nodes indicate the strength of their co-occurrence. Complementing this, Fig. 9 presents a trend analysis of keyword usage over recent years, indicating emerging areas of focus and shifts in scholarly attention. Based on the VOSviewer output, the keyword co-occurrence map (Fig. 8) is organized into seven distinct clusters, each representing a major research stream. **Erreur ! Source du renvoi introuvable.**Dominant themes are clearly centered on the integration of artificial intelligence with monetary and financial systems. The most frequently occurring keywords are "large language models" (17 occurrences) and "central bank communication" (9 occurrences), which underscores the growing convergence of advanced NLP techniques and monetary policy analysis. Closely related terms such as "machine learning" (9) and "monetary policy" (7) further emphasize the methodological and institutional focus of the literature.

Furthermore, cluster analysis (*Fig. 8*) shows that Cluster 1 (red), with terms like "bankruptcy prediction" and "financial ratios", represents traditional financial modeling and predictive analytics, with an average publication year around 2020. Cluster 2 (green) is the most recent and central theme during the average year 2025 (*Fig. 9*), focusing on sentiment analysis and its implications for central banking and financial stability. Cluster 7 (orange) focuses on GenAI technologies, including LLM-based tools, which mark the methodological frontier of the field. Moreover, *Fig. 9* reveals a clear chronological shift. Research has moved from traditional technical implementation around 2020 (Cluster 6, light blue: "C++" and "programming") toward cutting-edge applications of GenAI and LLMs in 2022 onwards (Cluster 2 and 7), reflecting a broader industry adoption curve.

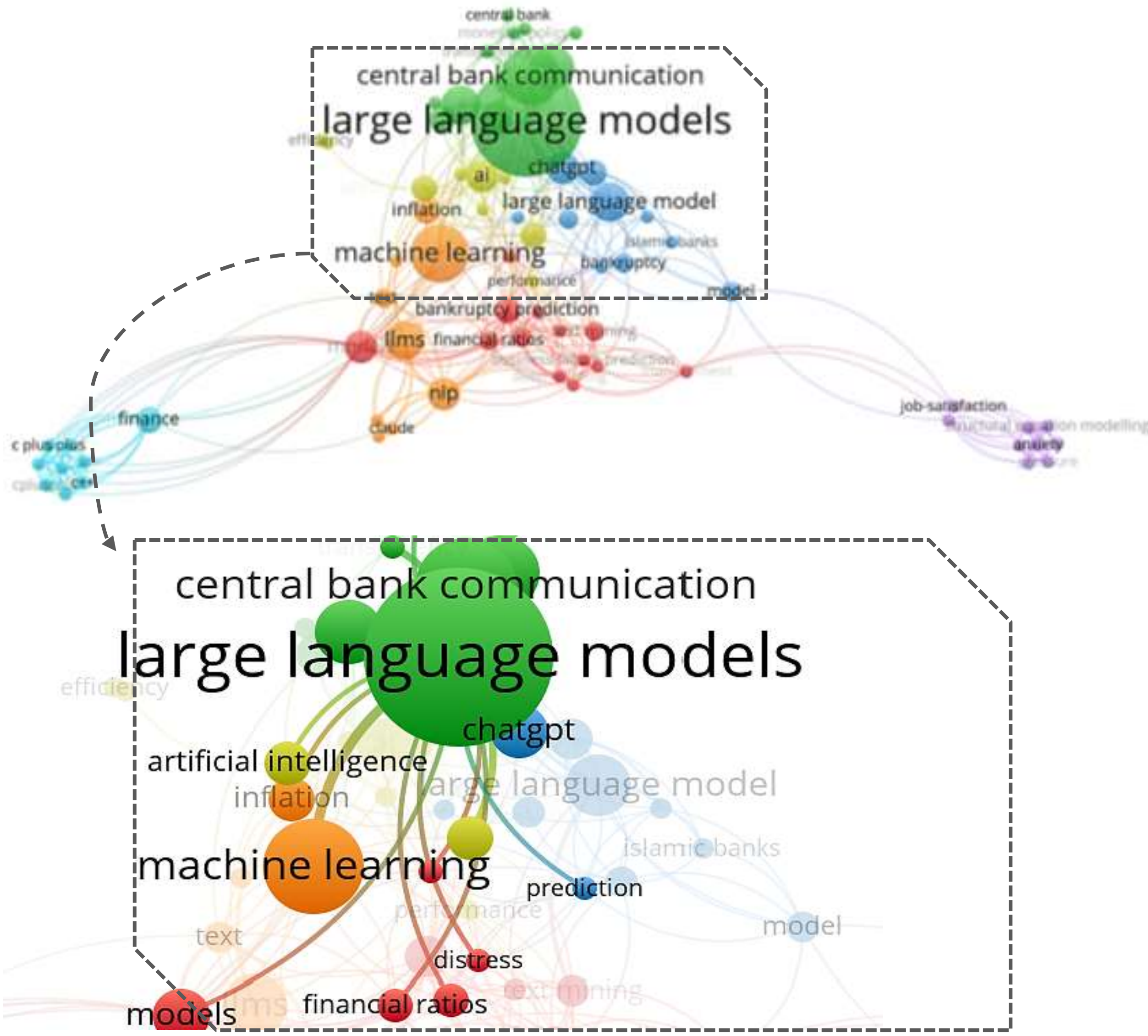


**Fig. 8** Co-occurrence map showing the frequency and connections of keywords in the reviewed studies

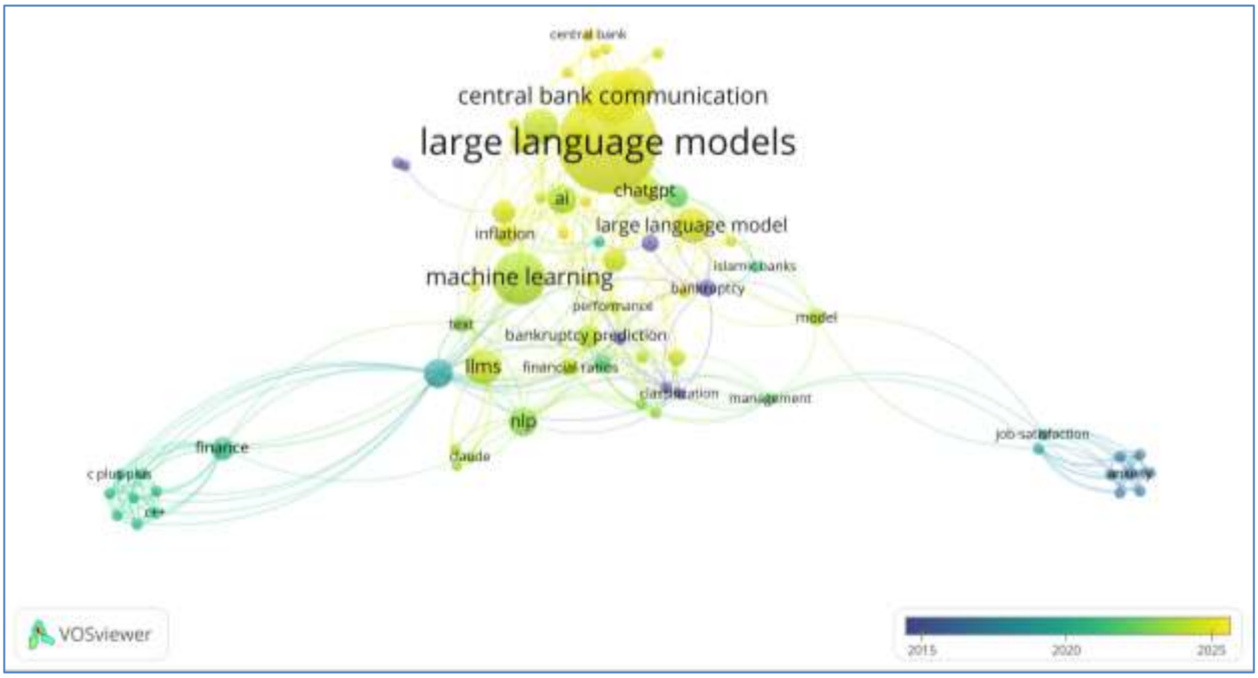


Fig. 9 **Temporal trends of frequently occurring keywords in the literature**

## 5. Large Language models in Banking management

### 5.1. Policy Interpretation and Sentiment

The shift from manual, dictionary-based methodologies to LLMs marked a clear improvement over traditional “bag-of-words” approaches that often fail to capture context. Pfeifer and Marohl (2023) showed that advanced models like CentralBankRoBERTa are necessary to distinguish between macroeconomic agents and identify emotional subtext in central bank communications. This transition toward contextual determination was further supported by Ardekani et al. (2024), who presented FinSentGPT as a universal engine capable of overcoming language barriers to predict sentiment across multi-language European Central Bank (ECB) datasets. The efficacy of these models was not merely theoretical; Alonso-Robisco and Carbo (2023) reported that sentiment analysis conducted using LLMs, specifically ChatGPT, closely aligns with assessments made by human experts when examining Central Bank Digital Currencies (CBDCs). Beyond identifying simple sentiment, researchers utilized LLMs to uncover the “shadow” mandates and structural shifts in central bank priorities. Bertsch et al. (2025) analyzed the Federal Reserve’s interpretation of its mandate and found that, although financial stability was not an explicit objective, it receives increased attention during periods of high debt-to-GDP. The institutional factors driving such communication were further probed by Leek and Bischl (2025). The authors utilized Google’s Gemini to demonstrate that increases in central bank independence (CBI) shift the focus of speeches from price stability toward broader financial pressures. To ensure such shifts in communication are transparently reflected, Ospina-Tejeiro and Romero (2025) applied Retrieval-Augmented Generation (RAG) to confirm that recent decision-making reforms in Colombia successfully increase monetary policy transparency. Similarly, Patterson (2025) introduced Amazon Nova to quantitatively evaluate the alignment between Federal Open Market Committee (FOMC) statements and their subsequent minutes. The study reported a consistent stability in central bank messaging over time.

The move toward context-aware modeling also allowed for more precise measurements of economic uncertainty. Ito et al. (2025) found that traditional keyword-based tracking fails to differentiate between types of uncertainty, whereas fine-tuned LLMs successfully captured the nuanced dynamics of Japan's monetary policy articles. Moreover, Ko et al. (2025) demonstrated that few-shot learning with general-purpose models (e.g., flan-T5 (Chung et al., 2022)) yields results comparable to domain-specific LLMs (e.g., finBERT), thereby highlighting the cost-effectiveness of general-purpose models. However, Baerg and Binder (2024) cautioned that the reliability of these automated detections of sentiment is highly dependent on the researcher's choices, as central bank communications remained subjectively difficult to understand despite technological advances. This sentiment was echoed by Kim et al. (2024), whose evaluation of Llama-3-70B and GPT-4 revealed that human expert performance remains significantly superior to all existing models.

Ultimately, the argument for integrating sentiment indices into economic frameworks was solidified by their predictive power regarding inflation and market behavior. Cho and Jung (2026) demonstrated that the tone of central bankers' speeches has asymmetric effects. Specifically, it significantly raised inflation expectations during economic expansions while remaining muted during downturns. Kanelis and Siklos (2025) found that the economic outlook shapes the content of ECB press conferences; however, sentiment regarding financial stability was often unaffected, indicating a selective use of narratives. To quantify these narratives, Fraccaroli et al. (2025) used LLMs to tag central bank speeches at the sentence level. This approach provided new insights into how institutions diagnosed the 2021 inflation surge. Simionescu and Nicula (2024) showed that incorporating sentiment indices from Inflation Reports into econometric models provides more accurate quarterly forecasts for Romania than traditional numerical methods. Furthermore, Davies and Hellings (2024) utilized LLMs to enhance the real-time tracking of trade-related inflation. Their approach aimed to provide central bankers and regulators with a policy-oriented tool for evaluating the impact of non-tariff barriers on national inflation. The analysis showed that EU-originating food prices increase significantly faster than domestic products.

The reach of these models even extended to the media, where Chuffart and Dell'Eva (2024) analyzed media coverage of the ECB by applying NLP to French newspaper articles. They used the Latent Dirichlet Allocation algorithm to group news into eleven thematic categories and constructed a specialized dictionary to assess media tone. Additionally, they employed the CamemBERT (Martin et al., 2020) LLM to calculate the probability of news being positive. These indices were ultimately integrated into a Vector Autoregression (VAR) model to evaluate the interaction between media narratives and central banking policy. Iadisernia and Camassa (2025) proved that GPT-4o can replicate professional forecasting panels for macroeconomic variables with remarkable accuracy. The model effectively employed synthetic personas to bridge the gap between AI and human expertise. Alonso et al. (2025) investigated public sentiment toward central banking policy, specifically the acceptance or rejection of CBDCs. Using ChatGPT 4.0 to generate 663 synthetic responses, the authors simulated how individuals with varying demographic backgrounds and levels of financial literacy might react to CBDC adoption. The study found that factors such as income, education, and financial experience are primary determinants of whether the public will support a central bank's digital liabilities. This work highlights the potential of GenAI as a low-cost tool for central banks to predict economic behavior regarding new policy instruments.

### 5.2. Risk Management and Financial Prediction

The banking management sector is fundamentally defined by the successful assessment, mitigation, and prediction of financial risk. This domain is undergoing a significant transformation through the integration of LLMs, GenAI, and advanced textual analytics. This technological integration allows banks to move beyond reliance on traditional quantitative metrics by leveraging rich, unstructured data to enhance predictive accuracy across multiple risk categories, including credit, bankruptcy, and systemic risk (Gupta et al., 2025; Lin et al., 2025; Wu et al., 2025). A primary argument in this shift is that qualitative narratives often contain early-warning signals of corporate distress that structured data fails to capture. For instance, Damiano et al. (2025) found that governance-related textual content in annual reports can be used as a proactive tool to detect corruption scandals before they become public. Similarly, Wu et al. (2025) demonstrated that integrating ChatGPT-refined loan assessments with conventional data significantly enhances default prediction. This integration also increased lending profitability over human-written texts. Borchert et al. (2023) further validated the necessity of unconventional data by showing that features extracted through transformers from company websites add substantial value to business failure prediction models.

This predictive evolution is particularly vital for small and medium-sized enterprises (SMEs) and mid-cap markets, where standard financial reporting is often sparse or delayed. Lin et al. (2025) suggested that LLMs like ChatGPT can bridge this gap by integrating social media and policy changes into credit rating models. These models provided satisfactory accuracy even for firms lacking public financial reports. Metzger et al. (2025) complemented this by examining how AI-generated financial diagnostics bolster SME management functions, offering real-time "red-flag" alerts for going concern risks. In the mid-cap sector, Korangi et al. (2023) utilized transformer-based architectures to predict default probability term structures. Their approach achieved superior performance by processing input data through a multi-channel architecture rather than traditional aggregation. Furthermore, Saleh and Semaan (2024) found that the financial sentiment of entrepreneurs is a viable predictor of their long-term survival and sustainability.

However, the rapid adoption of complex algorithms has faced some critical pushback regarding the necessity of such technical depth. Gupta et al. (2025) provided a notable counter-argument. They suggested that while nonlinear machine learning models can benefit from textual data, simple linear models using traditional financial ratios remain highly efficient. For certain bank failure predictions, this may make complex LLM algorithms redundant. Despite this perspective, a substantial portion of the literature emphasized that textual narratives capture contextual information that cannot be fully conveyed by numerical data alone. Rybinski (2021) validated this by ranking professional forecasters based on the predictive power of their narratives. The author found that optimal forecasting models almost always include an NLP index. Lis et al. (2024) also applied NLP to internal validation reports, grouping hundreds of findings into categories to identify systemic issues in credit risk model maintenance. More in this context, Sideras et al. (2024) claimed that incorporating auditor opinion texts is critical for assessing financial health, especially when utilizing data augmentation to address the severe class imbalance inherent in bankruptcy datasets.

The scope of financial prediction has also expanded to include emerging external risks, such as climate transition and geographic volatility. Lu and Wang (2025) developed a novel Stranded Asset Risk (SAR) index using a BERT-based model. They found that high SAR scores significantly reduce corporate leverage, particularly for fossil-energy-dependent firms. Moreno and Caminero (2024) utilized LLMs to calculate a Greenhouse Gas Disclosure Index (GHGDI) for European banks. Their study proved the effectiveness of AI in verifying the granularity of climate-related reporting. To address the black-box nature of such assessments, Yim et al. (2025) introduced an AI-powered framework using RAG to extract granular, transparent assessment factors for TCFD disclosures. Operational and systemic risks are also increasingly managed through agentic and sequence-based modeling. Skalski et al. (2023) found that generative pretraining on transaction sequences provides contextualized embeddings that outperform state-of-the-art methods in fraud detection and customer churn prediction. Huong et al. (2024) supported this approach by utilizing transaction-based graph learning to improve money laundering detection, which proved to be effective when combined with under-sampling techniques. On a systemic level, Stander (2024) constructed a news sentiment index using FinBERT and demonstrated its utility as a leading indicator of systemic risk that can inform IFRS 9 impairment quantification.

Ultimately, the goal of these predictive tools is to optimize human and organizational decision-making under uncertainty. Wang and Zhong (2025) used LLMs to show that innovation policy uncertainty has a dual effect: it encourages project initiation while simultaneously increasing the probability of bankruptcy. Furthermore, predictive success can be influenced by broader external factors, such as the geographic determinants of bankruptcy (Buehler et al., 2012), or the sociological shifts in exporter profiles driven by immigration (Tadesse & White, 2016). Beckmann and Hark (2024) provided a strategic caution, noting that the launch of tools like ChatGPT itself induces a negative stock market reaction for large, deposit-dependent banks. This response signaled a market perception of technological disruption.

### 5.3. Regulatory Technology (RegTech) and Compliance

Regulatory Technology (RegTech) represents a key area that affects banking management. In this domain, LLMs are used to address the growing complexity and scope of regulatory and reporting requirements, including emerging mandates related to ESG standards. Specifically, researchers are designing effective prompt engineering methods to guide LLMs in distilling dense regulatory texts (e.g., Basel III capital requirement regulations) into a concise mathematical framework that can subsequently be translated into actionable code. Cao and Feinstein (2024) conducted a case study evaluating the performance of several LLMs in this context. Their results showed that GPT-4 outperforms alternative models in mathematical reasoning, information extraction, and data collection tasks required for implementing Basel III capital adequacy requirements. This approach aimed to accelerate the deployment of regulatory mandates within financial reporting and risk management systems. It also reduced reliance on time-consuming manual specification and tagging, which often leads to incomplete regulatory coverage.

In an early study, Roychoudhury et al. (2018) discovered that traditional document-driven and expert-dependent ways of managing compliance are fundamentally flawed in the modern high-stakes regulatory environment. To address this limitation, the authors proposed a semi-automated framework in which domain experts author regulatory rules using a controlled natural language. This design helped bridge the gap between legal text and formal logic. The approach was validated through a case study on Money Market Statistical Reporting (MMSR) at a large European bank. The results demonstrated that converting legal English into a model form enables more rigorous and formal compliance checking than existing industry practices. This shift toward logical automation provided a foundational architecture for more advanced systems capable of supporting increasingly complex global mandates.

Beyond internal bank reporting, the effectiveness of RegTech must also extend to a supervisory level to ensure institutional consistency and fairness. Aarab (2025) noted that bank supervisors faced an increasingly complex task in ensuring that new measures are consistently aligned with a vast database of historical precedents. To address this challenge, the author introduced a novel Information Retrieval (IR) system designed to assist supervisors in drafting effective and consistent measures based on findings from on-site investigations. The system employed a sophisticated blend of lexical, semantic, and fuzzy set matching, allowing it to retrieve relevant historical measures with a high degree of precision. The efficacy of this system was further validated using Monte Carlo methodology. The results showed that hybrid semantic-lexical models significantly outperform standalone BERT-like or lexical models in maintaining regulatory integrity. Despite these advances, regulatory models proved to be undermined if human managerial discretion was not adequately considered. Soedarmono et al. (2017) provided a critical argumentative perspective by examining the implementation of the "Expected" Loan Loss Model (E-LLM) in Islamic banks. They demonstrated that loan loss provisioning remains stubbornly procyclical despite the implementation of the E-LLM, suggesting that the model itself does not automatically guarantee countercyclical stability. In addition, the authors demonstrated that Islamic banks often utilize loan loss provisions for discretionary capital management, which inflated reserves when bank capitalization was already in decline. Based on these findings, the study concluded that higher capitalization requirements can provide a more reliable safeguard than complex provisioning models, as managerial discretion can undermine the intended objectives of regulatory frameworks. In a comprehensive exploration of modern financial technology, Fan (2024) systematically analyzed the application of LLMs across the banking business value chain. The study evaluated extensive banking scenarios and identified technical dimensions where AI can be strategically implemented. The study highlights that while these models offered competitive advantages for banks, the industry must proactively address critical challenges such as data privacy, model interpretability, and algorithmic bias. The work provided strategic recommendations for banks to develop data ecosystems and compliance management frameworks for technological transformation. Moayeri et al. (2024) utilized institutional data from the World Bank to evaluate the reliability of AI models in recalling global facts. By creating "WorldBench" (i.e., a framework based on 11 individual World Bank indicators), the authors identified significant geographic and income-level biases in LLMs. A critical finding for banking and financial integrity was the detection of "citation hallucinations", where AI models provided false statistics while incorrectly claiming the World Bank as the source. This study warned that models are 1.5 times more likely to make errors regarding countries in Sub-Saharan Africa than those in North America.

### 5.4. Operational Efficiency and Process Management

Moving from the externally focused applications of RegTech, this review shifts toward the internal transformation of Operational Efficiency and Process Management. In this area, banking management has leveraged advanced computational and engineering techniques to streamline core activities, manage large transaction volumes, and improve internal productivity. In the contemporary banking environment, improving internal business processes and managing large volumes of textual data have become paramount. As organizations frequently manage numerous process models, automating the identification of process weaknesses has emerged as a key objective for business process improvement.

Bergener et al. (2015) claimed that the traditional manual analysis of process model landscapes is prohibitively time-consuming and resource-intensive. To address this issue, the authors utilized semantic pattern matching to automatically detect process weaknesses in a German bank. Further, the study reported that the proposed approach significantly supports business analysts by identifying improvement potentials more efficiently. This push toward automation was further advanced by Ang et al. (2025), who proposed a modular agentic framework, termed TS-Agent, designed to automate financial time-series modeling. They demonstrated that the use of reflective feedback and structured knowledge banks enables agentic systems to outperform traditional Automated Machine Learning (AutoML) frameworks.

A critical component of operational efficiency involved the ability of non-technical staff to interact with complex data. J. Wang et al. (2025) developed an LLM-enhanced text mining workflow that allows employees to analyze supply chain finance data without programming skills. The model used internet-sourced text information from bidding websites and financial statements in the Chinese new energy bus market. The findings indicated that the model, validated on a case study, is more practical and efficient than traditional manual auditing.

Similarly, Zhang et al. (2024) contended that generic Natural Language to SQL (NL2SQL) technologies are insufficient for the complexities of intra-enterprise financial big data. They implemented a system combining RAG and intelligent body feedback, which successfully improved SQL generation accuracy in financial environments from 54% to 70%. This emphasis on accurate data interaction was further supported by Kim et al. (2025), who introduced KoBankIR, a benchmark for banking information retrieval. The authors found that existing retrieval models are fundamentally incapable of handling the multi-document queries common in banking. As a result, the authors emphasized the need for systematic, domain-specific benchmarks to ensure reliable AI services.

Operational efficiency also extended to the automated generation of institutional knowledge and evaluation resources. Sayed et al. (2024) showed that RAG-based architectures allow for the creation of diverse and contextually relevant question banks. They reported that combining LLMs with domain-specific knowledge bases surpasses the quality and relevance of template-based generation. In the context of technical education for banking IT sectors, Sychev and Shashkov (2025) developed an automated generator that uses language-independent "meaning trees" to produce over 1.4 million programming problems. The study claimed that the proposed methodology not only ensured full correctness but also maintained a natural look that state-of-the-art randomized methods often failed to achieve. Furthermore, researchers re-evaluated the underlying economic factors that dictated technical efficiency. Tsionas and Mamatzakis (2017) challenged the standard assumption that inputs can freely adjust in efficiency models. The authors employed non-parametric local linear maximum likelihood to measure technical efficiency in global banking. They found that bank profit efficiency is directly constrained by adjustment costs in variable inputs. Their analysis of global banking revealed that personnel expenses constitute the highest adjustment costs in advanced countries, suggesting that efficiency gains are often hindered by labor-related rigidities. Additionally, Prakash et al. (2021) explored the potential for systemic efficiency through distributed intelligent agents. They proposed a system where on-board agents detected traffic violations and communicated directly with banking servers to deduct fines. This automated, corruption-free process eliminated the inefficiencies associated with manual detection and collection.

### 5.5. Customer-Facing Applications and Advisory Services

A key application for banking managers is employing LLMs for immediate customer service, specifically for accurately detecting a customer's financial intent (Loukas et al., 2023; Srivastava, 2024). In this context, Hens and Nordlie (2025) evaluated the capacity of OpenAI's ChatGPT-4 and Google's Gemini to replicate human bank experts in investor risk profiling. They found that AI-generated profiles are statistically comparable to expert assessments in approximately half of the cases. However, the models showed clear limitations in providing high-quality and useful explanations for their conclusions. This shortcoming in advisory reliability was also highlighted in the study of Lakkaraju et al. (2023), which compared LLM-based chatbots with rule-based systems in the personal finance domain. The authors reported that general-purpose LLMs are prone to retrieval errors and lack domain-specific reasoning capabilities for banking products, often providing inconsistent responses across different user groups compared to their rule-based counterparts.

Regarding operational intent identification, Dündar et al. (2020) implemented a large-scale system for Turkish banking messages. The results showed that contextual word embeddings like BERT and ELMo, when trained on 6,453 annotated messages, are essential for identifying 148 distinct customer intents. The models significantly boosted performance over traditional classification tasks. Srivastava (2024) further supported this effort by investigating the performance-cost trade-offs of using various instruction-tuned LLMs on the Banking77 dataset. The study demonstrated that smaller, open-source models, specifically when enhanced by RAG, can outperform larger, closed-source models. These models offered a substantially improved performance-cost ratio, making them more suitable for budget-limited organizations. Similar conclusions were drawn by Loukas et al. (2023), who investigated resource-limited text classification in banking. The authors noted that traditional NLP models require substantial amounts of labeled examples, which often proved impractical in data-limited banking domains. In contrast, few-shot learning in LLMs demonstrated effective performance in reducing operational costs while managing the scarcity of labeled data.

The strategic use of customer-generated data for service improvement was also examined. Kim and Ryu (2025) introduced a novel Importance-Performance Analysis (IPA) method using online reviews for Korean mobile banking apps. They utilized Latent Dirichlet Allocation (LDA) topic modeling and KoBERT-based sentiment analysis to evaluate critical service attributes. The findings showed that the proposed automated approach preserves measurement accuracy while offering cost-effective and intuitive metric assessments for the banking industry. Xi et al. (2024) focused on enhancing the personalization of digital advice by developing a memory-enhanced sequential conversational recommender system (CRS), referred to as MemoCRS. The system was trained on large-scale Chinese and English conversational datasets to address the "cold-start" problem and reduce noise from historical dialogue sessions. Experimental results showed that MemoCRS delivers more precise and tailored financial recommendations. Dwivedi et al. (2023) synthesized perspectives from 43 multidisciplinary experts to assess the impact of generative conversational AI on various sectors, including the banking industry. The authors highlighted that banking is one of the primary sectors expected to achieve significant productivity gains through the implementation of technologies like ChatGPT. While these tools can enhance banking management and marketing activities, the study cautioned that they also introduce systemic risks. These include potential disruptions to traditional banking practices, threats to data privacy, and the security consequences of algorithmic bias and misinformation.

## 6. Conclusion and Future Research Directions

### 6.1. Conclusion

This article presents a systematic state-of-the-art review and synthesis of 91 papers centered exclusively on banking management within the era of GenAI. By analyzing publications spanning from 2015 to 2025, and including forthcoming research, a structured taxonomy has been developed across five primary domains: Policy Interpretation and Sentiment, Risk Management and Financial Prediction, RegTech and Compliance, Operational Efficiency and Process Management, and Customer-Facing Applications and Advisory Services. The findings confirm a massive surge in scholarly attention beginning in 2023, with Risk Management and Financial Prediction emerging as the most significant research area, accounting for nearly 29% of the literature.

This review highlights an important shift within the industry, moving from traditional quantitative modeling toward context-aware tools capable of interpreting complex central bank mandates and detecting early-warning signals of corporate distress within unstructured narratives. While specialized financial models and hybrid pipelines, such as RAG and modular agentic frameworks, are increasingly prevalent, the field continues to face critical barriers related to data privacy and model interpretability. Furthermore, evidence suggests that existing models often struggle with complex, domain-specific multi-document queries and exhibit significant geographic disparities in factual accuracy. The author hopes that this systematic review will serve as a foundational reference for researchers and practitioners aiming to build transparent and field-compatible financial modeling systems in the evolving banking landscape.

### 6.2. Limitations and Future Directions

Based on the review presented in this article, the following are research avenues that may be addressed in future works to close existing gaps:

(1) **Development of Federated Learning Frameworks for Non-Public SME Credit Scoring.** While Lin et al. (2025) discussed using ChatGPT for rating SMEs without public reports, their testing was limited by small sample sizes. Future research could investigate the implementation of federated learning to allow multiple banks to train shared LLMs on private, sensitive SME data without exposing personally identifiable information.

(2) **Autonomous "Self-Healing" Regulatory Compliance Systems.** Current frameworks distill regulatory texts into mathematical logic for checking (Cao & Feinstein, 2024; Roychoudhury et al., 2018). A novel research path would be the development of autonomous agents capable of "self-healing," where a change in a regulation (e.g., Basel III updates) triggers an automated, end-to-end update of the bank's internal software code and user interfaces without human developer intervention.

(3) **Geographically Neutral Knowledge Distillation and Bias Correction.** To address the geographic disparities identified by Moayeri et al. (2024), research could focus on "Regional Knowledge Injection" techniques. This would involve developing training pipelines that prioritize the distillation of localized, high-fidelity economic datasets from emerging markets into the core reasoning layers of global financial LLMs to prevent the systematic disadvantage of Sub-Saharan or low-income regions in automated advisory services.

## Data Availability Statement

For transparency and reproducibility, the dataset used in this study (91 research articles) is provided in a ".txt" file. The corresponding VOSviewer map files related to author collaboration and keyword co-occurrence are also included as JSON files.

## References


- Aarab, I. (2025). LLM-based IR-system for bank supervisors. *Knowledge-Based Systems*, *310*. https://doi.org/10.1016/j.knosys.2024.112914
- Aldasoro, I., Gambacorta, L., Korinek, A., Shreeti, V., & Stein, M. (2025). Intelligent financial system: How AI is transforming finance. *Journal of Financial Stability*, *81*. https://doi.org/10.1016/j.jfs.2025.101472
- Ali, H., Zafar, M. B., & Aysan, A. F. (2025). Generative AI in finance: Replicability, methodological contingencies, and future research directions. *Finance Research Letters*, *86*(F). https://doi.org/10.1016/j.frl.2025.108797
- Alonso-Robisco, A., & Carbó, J. M. (2023). Analysis of CBDC narrative by central banks using large language models. *Finance Research Letters*, *58*(C). https://doi.org/10.1016/j.frl.2023.104643
- Alonso, S. L. N., Ozili, P. K., Hernández, B. M. S., & Pacheco, L. M. (2025). Evaluating the acceptance of CBDCs: experimental research with artificial intelligence (AI) generated synthetic response. *Quantitative Finance and Economics*, *9*(1), 242–273. https://doi.org/10.3934/QFE.2025008
- Ang, Y., Bao, Y., Jiang, L., Tao, J., Tung, A. K. H., Szpurch, L., & Ni, H. (2025). Structured Agentic Workflows for Financial Time-Series Modelling with LLMs and Reflective Feedback. *ICAIF 2025 - 6th ACM International Conference on AI in Finance*, 924–932. https://doi.org/10.1145/3768292.3771251
- Ardekani, A. M., Bertz, J., Bryce, C., Dowling, M., & Long, S. (2024). FinSentGPT: A universal financial sentiment engine? *International Review of Financial Analysis*, *94*. https://doi.org/10.1016/j.irfa.2024.103291
- Bae, J., Yu Hung, C., & van Lent, L. (2023). Mobilizing Text As Data. *European Accounting Review*, *32*(5), 1085–1106. https://doi.org/10.1080/09638180.2023.2218423
- Baerg, N., & Binder, C. (2024). Automated Detection of Emotion in Central Bank Communication: a Warning. *National Institute Economic Review*, *269*(SI), 82–91. https://doi.org/10.1017/nie.2024.31
- Beckmann, L., & Hark, P. F. (2024). ChatGPT and the banking business: Insights from the US stock market on potential implications for banks. *Finance Research Letters*, *63*. https://doi.org/10.1016/j.frl.2024.105237
- Bengio, Y., Ducharme, R., Vincent, P., Jauvin, C., Ca, J. U., Kandola, J., Hofmann, T., Poggio, T., & Shawe-Taylor, J. (2003). A Neural Probabilistic Language Model. *Journal of Machine Learning Research*, *3*, 1137–1155.
- Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *Npj Digital Medicine 2020 3:1*, *3*(1), 1–8. https://doi.org/10.1038/s41746-020-00324-0
- Bergener, P., Delfmann, P., Weiss, B., & Winkelmann, A. (2015). Detecting potential weaknesses in business processes: An exploration of semantic pattern matching in process models. *Business Process Management Journal*, *21*(1), 25–54. https://doi.org/10.1108/BPMJ-07-

2013-0103
- Bertsch, C., Hull, I., Lumsdaine, R. L., & Zhang, X. (2025). Central bank mandates and monetary policy stances: Through the lens of Federal Reserve speeches. *Journal of Econometrics*, *249*(A). https://doi.org/10.1016/j.jeconom.2025.105948
- Borchert, P., Coussement, K., De Caigny, A., & De Weerdt, J. (2023). Extending business failure prediction models with textual website content using deep learning. *European Journal of Operational Research*, *306*(1), 348–357. https://doi.org/10.1016/j.ejor.2022.06.060
- 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., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *2020-December*. https://arxiv.org/abs/2005.14165v4
- Buehler, S., Kaiser, C., & Jaeger, F. (2012). The geographic determinants of bankruptcy: Evidence from Switzerland. *Small Business Economics*, *39*(1), 231–251. https://doi.org/10.1007/s11187-010-9301-8
- Cao, Z., & Feinstein, Z. (2024). Large Language Model in Financial Regulatory Interpretation. *2024 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics, CIFEr 2024*. https://doi.org/10.1109/CIFER62890.2024.10772991
- Cho, D., & Jung, J. (2026). Mind the tone: Responses of inflation expectations to central bankers' speeches. *Journal of International Money and Finance*, *160*. https://doi.org/10.1016/j.jimonfin.2025.103452
- Chuffart, T., & Dell'Eva, C. (2024). Media Coverage of the ECB: a Textual Analysis. *Revue d'économie Politique*, *Vol. 134*(6), 923–945. https://doi.org/10.3917/redp.346.0923
- Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tai, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S. S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Castro-Ros, A., Pellat, M., Robinson, K., … Wei, J. (2022). Scaling Instruction-Finetuned Language Models. *Journal of Machine Learning Research*, *25*. https://arxiv.org/pdf/2210.11416
- Coelho E Silva, L., Fonseca, G. D. F., & Castro, P. A. L. (2024). Transformers and attention-based networks in quantitative trading: a comprehensive survey. *ICAIF 2024 - 5th ACM International Conference on AI in Finance*, 822–830. https://doi.org/10.1145/3677052.3698684
- Damiano, R., Polizzi, S., Scannella, E., & Valenza, G. (2025). Corruption Detection Through Textual Analysis: Evidence From Eurozone Banks. *Business Ethics, the Environment and Responsibility*. https://doi.org/10.1111/beer.12824
- Davies, R., & Hellings, J. (2024). Britain'S Trade Challenge: Tracking the Costs in Real Time. *National Institute Economic Review*, *268*, 15–29. https://doi.org/10.1017/nie.2024.15
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Naacl-Hlt 2019*, *Mlm*, 4171–4186. https://aclanthology.org/N19-1423.pdf
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 4171–4186. https://arxiv.org/abs/1810.04805v2
- Dong, M. M., Stratopoulos, T. C., & Wang, V. X. (2024). A scoping review of ChatGPT research in accounting and finance. *International Journal of Accounting Information Systems*, *55*. https://doi.org/10.1016/j.accinf.2024.100715
- Dündar, E. B., Kiliç, O. F., Çekiç, T., Manav, Y., & Deniz, O. (2020). Large scale intent

detection in turkish short sentences with contextual word embeddings. *IC3K 2020 - Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, *1*, 187–192. https://doi.org/10.5220/0010108301870192
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*. https://doi.org/10.1016/j.ijinfomgt.2023.102642
- Eisfeldt, A. L., & Schubert, G. (2025). Generative AI and Finance. *Annual Review of Financial Economics*, *17*(1), 363–393. https://doi.org/10.1146/annurev-financial-112923-020503
- Elsevier. (2020). *Mendeley Reference Manager* (v1.19.8). https://www.mendeley.com/download-reference-manager/windows
- Fan, M. (2024). LLMs in Banking: Applications Challenges and Approaches. *Proceedings of International Conference on Digital Economy, Blockchain and Artificial Intelligence, DEBAI 2024*, 314–321. https://doi.org/10.1145/3700058.3700107
- Feng, Z., Hu, G., Li, B., & Wang, J. (2025). Unleashing the power of ChatGPT in finance research: opportunities and challenges. *Financial Innovation*, *11*(1). https://doi.org/10.1186/s40854-025-00770-3
- Fraccaroli, N., Arel-Bundock, V., & Blyth, M. (2025). What do central bankers talk about when they talk about inflation? The rise and fall of inflation narratives. *New Political Economy*, *30*(5), 713–728. https://doi.org/10.1080/13563467.2025.2504392
- Gupta, A., Lu, C., Simaan, M., & Zaki, M. J. (2025). When Positive Sentiment is not so Positive: Textual Analytics and Bank Failures. *Computational Economics*. https://doi.org/10.1007/s10614-025-10969-2
- Haidar, A., & Abbass, A. (2025). Navigating the Frontier of Finance: A Scoping Review of Generative AI Applications and Implications. *Generative Artificial Intelligence in Finance: Large Language Models, Interfaces, and Industry Use Cases to Transform Accounting and Finance Processes*, 215–252. https://doi.org/10.1002/9781394271078.ch12
- Hens, T., & Nordlie, T. (2025). How good are LLMs in risk profiling? *Finance Research Letters*, *85*(D). https://doi.org/10.1016/j.frl.2025.108102
- Huong, H., Nguyen, X., Dang, T. K., & Tran-Truong, P. T. (2024). Money Laundering Detection Using A Transaction-Based Graph Learning Approach. *Proceedings of the 2024 18th International Conference on Ubiquitous Information Management and Communication, IMCOM 2024*. https://doi.org/10.1109/IMCOM60618.2024.10418307
- Iadisernia, G., & Camassa, C. (2025). Prompting for Policy: Forecasting Macroeconomic Scenarios with Synthetic LLM Personas. *ICAIF 2025 - 6th ACM International Conference on AI in Finance*, 335–343. https://doi.org/10.1145/3768292.3770385
- Ito, A., Sato, M., & Ota, R. (2025). A novel content-based approach to measuring monetary policy uncertainty using fine-tuned LLMs. *Finance Research Letters*, *75*. https://doi.org/10.1016/j.frl.2025.106832
- Jelinek, F. (1976). Continuous Speech Recognition by Statistical Methods. *Proceedings of the IEEE*, *64*(4), 532–556. https://doi.org/10.1109/PROC.1976.10159
- Kanelis, D., & Siklos, P. L. (2025). The ECB press conference statement: deriving a new sentiment indicator for the euro area. *International Journal of Finance and Economics*, *30*(1), 652–664. https://doi.org/10.1002/ijfe.2940
- Kim, H., Yoo, Y., & Kwak, Y. (2025). Query Generation Pipeline with Enhanced Answerability

Assessment for Financial Information Retrieval. *ICAIF 2025 - 6th ACM International Conference on AI in Finance*, 141–149. https://doi.org/10.1145/3768292.3770354
– Kim, S., & Ryu, M. H. (2025). Online review based IPA and IPCA: the case of Korean mobile banking apps. *International Journal of Bank Marketing*, *43*(4), 731–756. https://doi.org/10.1108/IJBM-03-2024-0136
– Kim, W., Spörer, J., Lee, C. L., & Handschuh, S. (2024). Is Small Really Beautiful for Central Bank Communication? Evaluating Language Models for Finance: Llama-3-70B, GPT-4, FinBERT-FOMC, FinBERT, and VADER. *ICAIF 2024 - 5th ACM International Conference on AI in Finance*, 626–633. https://doi.org/10.1145/3677052.3698675
– Ko, E., Dmonte, A., & Zampieri, M. (2025). Comparison of an affine term structure model with Fed chair speeches in large language models. *Finance Research Letters*, *78*. https://doi.org/10.1016/j.frl.2025.107114
– Kong, Y., Nie, Y., Dong, X., Mulvey, J. M., Poor, H. V., Wen, Q., & Zohren, S. (2024). Large Language Models for Financial and Investment Management: Applications and Benchmarks. *Journal of Portfolio Management*, *51*(2), 162–210. https://doi.org/10.3905/jpm.2024.1.645
– Korangi, K., Mues, C., & Bravo, C. (2023). A transformer-based model for default prediction in mid-cap corporate markets. *European Journal of Operational Research*, *308*(1), 306–320. https://doi.org/10.1016/j.ejor.2022.10.032
– Lakkaraju, K., Jones, S. E., Vuruma, S. K. R., Pallagani, V., Muppasani, B. C., & Srivastava, B. (2023). LLMs for Financial Advisement: A Fairness and Efficacy Study in Personal Decision Making. *ICAIF 2023 - 4th ACM International Conference on AI in Finance*, 100–107. https://doi.org/10.1145/3604237.3626867
– Lee, D. K. C., Guan, C., Yu, Y., & Ding, Q. (2024). A Comprehensive Review of Generative AI in Finance. *FinTech*, *3*(3), 460–478. https://doi.org/10.3390/fintech3030025
– Leek, L., & Bischl, S. (2025). How central bank independence shapes monetary policy communication: A Large Language Model application. *European Journal of Political Economy*, *87*. https://doi.org/10.1016/j.ejpoleco.2025.102668
– Lengyel, P., Pancsira, J., & Füzesi, I. (2025). *Machine Learning Integration in Cryptocurrency Trading: A Systematic Review of Fintech Implications*. https://doi.org/10.21203/RS.3.RS-7411123/V1
– Li, Y., Wang, S., Ding, H., & Chen, H. (2023). Large Language Models in Finance: A Survey. *ICAIF 2023 - 4th ACM International Conference on AI in Finance*, 374–382. https://doi.org/10.1145/3604237.3626869
– Lin, J., Lai, S., Yu, H., Liang, R., & Yen, J. (2025). ChatGPT based credit rating and default forecasting. *Journal of Data, Information and Management*, *7*(1), 69–92. https://doi.org/10.1007/s42488-025-00143-6
– Lis, S., Kubkowski, M., Borkowska, O., Serwa, D., & Kurpanik, J. (2024). Analyzing credit risk model problems through natural language processing-based clustering and machine learning: insights from validation reports. *Journal of Risk Model Validation*, *18*(2), 59–86. https://doi.org/10.21314/JRMV.2024.006
– Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V., & Allen, P. G. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. https://arxiv.org/pdf/1907.11692
– Loukas, L., Stogiannidis, I., Diamantopoulos, O., Malakasiotis, P., & Vassos, S. (2023). Making LLMs Worth Every Penny: Resource-Limited Text Classification in Banking. *ICAIF 2023 - 4th ACM International Conference on AI in Finance*, 392–400. https://doi.org/10.1145/3604237.3626891

- Lu, H., & Wang, X. (2025). Stranded asset risk and corporate capital structure: Evidence from China's low-carbon transition. *Research in International Business and Finance*, *80*. https://doi.org/10.1016/j.ribaf.2025.103144
- Mahendran, M. B., Gokul, A. K., Lakshmi, P., & Pavithra, S. (2025). Comparative Advances in Financial Sentiment Analysis:A Review of BERT,FinBert, and Large Language Models. *3rd International Conference on Intelligent Data Communication Technologies and Internet of Things, IDCIoT 2025*, 39–45. https://doi.org/10.1109/IDCIOT64235.2025.10914764
- Martin, L., Muller, B., Suárez, P. J. O., Dupont, Y., Romary, L., de la Clergerie, É. V., Seddah, D., & Sagot, B. (2020). CamemBERT: a Tasty French Language Model. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 7203–7219. https://doi.org/10.18653/v1/2020.acl-main.645
- Mohamed, Maha Mohamed Alsebai, Mohamed Djafar Henni, and Nema Amin Alsayed Sorour. 2026. 'Integrating Digital and AI-Driven Productivity into National Accounts: A Systemic Analysis of Economic Impacts in Emerging and Advanced Economies', *Sustainability*, 18: 878.
- Metzger, M., O'Reilly, S., & Mac an Bhaird, C. (2025). Generative artificial intelligence augmenting SME financial management. *Technovation*, *147*. https://doi.org/10.1016/j.technovation.2025.103313
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*. https://arxiv.org/abs/1301.3781v3
- Mo, H., & Ouyang, S. (2025). (Generative) AI in Financial Economics. *Journal of Chinese Economic and Business Studies*, *23*(4), 509–587. https://doi.org/10.1080/14765284.2025.2569006
- Moayeri, M., Tabassi, E., & Feizi, S. (2024). WorldBench: Quantifying Geographic Disparities in LLM Factual Recall. *2024 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2024*, 1211–1228. https://doi.org/10.1145/3630106.3658967
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/J.IJSU.2010.02.007
- Moreno, A. I., & Caminero, T. (2024). Assessing climate-related disclosures of European banks through text mining. *Review of World Economics*. https://doi.org/10.1007/s10290-024-00531-x
- Muggleton, S. (2014). Alan turing and the development of artificial intelligence. *AI Communications*, *27*(1), 3–10. https://doi.org/10.3233/AIC-130579
- OpenAI. (2024). *ChatGPT*. [Large Language Model]. https://chat.openai.com/chat
- Ospina-Tejeiro, J. J., & Romero, J. V. (2025). Monetary policy transparency in Colombia. *Latin American Journal of Central Banking*. https://doi.org/10.1016/j.latcb.2025.100185
- Patterson, A. (2025). Assessing the Alignment of FOMC Statements with Minutes using Large Language Models. *Issues in Information Systems*, *26*(2), 96–108. https://doi.org/10.48009/2_iis_108
- Peng, X., Han, C., Ouyang, F., & Liu, Z. (2020). Topic tracking model for analyzing student-generated posts in SPOC discussion forums. *International Journal of Educational Technology in Higher Education*, *17*(1), 1–22. https://doi.org/10.1186/S41239-020-00211-4/TABLES/6
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language*

*Processing, Proceedings of the Conference*, 1532–1543. https://doi.org/10.3115/V1/D14-1162
– Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. *NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 2227–2237. https://doi.org/10.18653/v1/n18-1202
– Pfeifer, M., & Marohl, V. P. (2023). CentralBankRoBERTa: A fine-tuned large language model for central bank communications. *Journal of Finance and Data Science*, *9*. https://doi.org/10.1016/j.jfds.2023.100114
– Prakash, P., Shreeya, Y. C., Srivastava, A., Chaitanya Priya, K. J., & Sharma, A. (2021). A Conceptual Model Simulation to Detect and Report City Traffic Violations using Distributed Intelligent Agents. *2021 6th International Conference for Convergence in Technology, I2CT 2021*. https://doi.org/10.1109/I2CT51068.2021.9417881
– Radford, A., & Narasimhan, K. (2018). *Improving Language Understanding by Generative Pre-Training*.
– Raliphada, P., Olusanya, M., & Olukanmi, S. (2025). *Transformer-based NLP Approaches for Credit Risk Prediction: A Systematic Review*. https://www.researchsquare.com/article/rs-7707021/v1
– Roychoudhury, S., Sunkle, S., Choudhary, N., Kholkar, D., & Kulkarni, V. (2018). A case study on modeling and validating financial regulations using (semi-) automated compliance framework. *Lecture Notes in Business Information Processing*, *335*, 288–302. https://doi.org/10.1007/978-3-030-02302-7_18
– Rybinski, K. (2021). Ranking professional forecasters by the predictive power of their narratives. *International Journal of Forecasting*, *37*(1), 186–204. https://doi.org/10.1016/j.ijforecast.2020.04.003
– Saleh, L., & Semaan, S. (2024). The Potential of AI in Performing Financial Sentiment Analysis for Predicting Entrepreneur Survival. *Administrative Sciences*, *14*(9). https://doi.org/10.3390/admsci14090220
– Saxena, A., Verma, S., & Mahajan, J. (2024). Generative AI in banking financial services and insurance: A guide to use cases, approaches, and insights. *Generative AI in Banking Financial Services and Insurance: A Guide to Use Cases, Approaches, and Insights*, 1–346. https://doi.org/10.1007/979-8-8688-0559-2
– Sayed, A., Kanojia, M., & Nabajja, S. (2024). Advanced Subjective Question Bank Generation Using Retrieval Augmented Generation Architecture. *International Journal of Computer Information Systems and Industrial Management Applications*, *16*(3), 264–277. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85201672275&partnerID=40&md5=03ece3f3ff81a9dcd748794841f45b01
– Sideras, A., Bougiatiotis, K., Zavitsanos, E., Paliouras, G., & Vouros, G. (2024). Bankruptcy Prediction: Data Augmentation, LLMs and the Need for Auditor's Opinion. *ICAIF 2024 - 5th ACM International Conference on AI in Finance*, 453–460. https://doi.org/10.1145/3677052.3698627
– Simionescu, M., & Nicula, A.-S. (2024). Sentiment Analysis as an Innovation in Inflation Forecasting in Romania. *Marketing and Management of Innovations*, *15*(2), 13–25. https://doi.org/10.21272/mmi.2024.2-02
– Skalski, P., Sutton, D., Burrell, S., Perez, I., & Wong, J. (2023). Towards a Foundation Purchasing Model: Pretrained Generative Autoregression on Transaction Sequences. *ICAIF 2023 - 4th ACM International Conference on AI in Finance*, 141–149.

https://doi.org/10.1145/3604237.3626850
- Soedarmono, W., Pramono, S. E., & Tarazi, A. (2017). The procyclicality of loan loss provisions in Islamic banks. *Research in International Business and Finance*, *39*(B), 911–919. https://doi.org/10.1016/j.ribaf.2016.05.003
- Srivastava, V. (2024). Lending an Ear: How LLMs Hear Your Banking Intentions. *ICAIF 2024 - 5th ACM International Conference on AI in Finance*, 301–309. https://doi.org/10.1145/3677052.3698608
- Stander, Y. S. (2024). A News Sentiment Index to Inform International Financial Reporting Standard 9 Impairments. *Journal of Risk and Financial Management*, *17*(7). https://doi.org/10.3390/jrfm17070282
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. *Advances in Neural Information Processing Systems*, *4*(January), 3104–3112. https://arxiv.org/abs/1409.3215v3
- Sychev, O., & Shashkov, D. (2025). Mass Generation of Programming Learning Problems from Public Code Repositories. *Big Data and Cognitive Computing*, *9*(3). https://doi.org/10.3390/bdcc9030057
- Tadesse, B., & White, R. (2016). Do immigrants affect the profile of U.S. exporters? *Applied Economics*, *48*(19), 1743–1758. https://doi.org/10.1080/00036846.2015.1105930
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models*. https://arxiv.org/pdf/2302.13971
- Tsionas, E. G., & Mamatzakis, E. C. (2017). Adjustment costs in the technical efficiency: An application to global banking. *European Journal of Operational Research*, *256*(2), 640–649. https://doi.org/10.1016/j.ejor.2016.06.037
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 435–460. https://doi.org/10.1007/978-1-4020-6710-5_3
- van Eck, N. J., & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics 2009 84:2*, *84*(2), 523–538. https://doi.org/10.1007/S11192-009-0146-3
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, *2017-December*, 5999–6009. https://arxiv.org/abs/1706.03762v7
- Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. R. (2018). GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. *EMNLP 2018 - 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP, Proceedings of the 1st Workshop*, 353–355. https://doi.org/10.18653/v1/w18-5446
- Wang, J., Liu, G., Cheng, Y., Xu, X., & Li, Z. (2025). Leveraging Internet-Sourced Text Data for Financial Analytics in Supply Chain Finance: A Large Language Model-Enhanced Text Mining Workflow. *IEEE Transactions on Engineering Management*, *72*, 1924–1938. https://doi.org/10.1109/TEM.2025.3567302
- Wang, K. (Wei), & Zhong, C. (Xiaoqian). (2025). The dark side of innovation policy uncertainty. *Finance Research Letters*, *85*. https://doi.org/10.1016/j.frl.2025.108066
- Wang, Y., Li, H., & Liao, L. (2020). DEM Construction Method for Slopes Using Three-Dimensional Point Cloud Data Based on Moving Least Square Theory. *JOURNAL OF SURVEYING ENGINEERING*, *146*(3). https://doi.org/10.1061/(ASCE)SU.1943-5428.0000320
- Weizenbaum, J. (1966). ELIZA-A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

https://doi.org/10.1145/365153.365168/ASSET/52A33E60-61E1-440C-8557-F51EDCC5D9BF/ASSETS/365153.365168.FP.PNG

- Wu, Z., Dong, Y., Li, Y., & Shi, B. (2025). Unleashing the power of text for credit default prediction: Comparing human-written and generative AI-refined texts. *European Journal of Operational Research*, *326*(3), 691–706. https://doi.org/10.1016/j.ejor.2025.04.032
- Xi, Y., Liu, W., Lin, J., Chen, B., Tang, R., Zhang, W., & Yu, Y. (2024). MemoCRS: Memory-enhanced Sequential Conversational Recommender Systems with Large Language Models. *International Conference on Information and Knowledge Management, Proceedings*, 2585–2595. https://doi.org/10.1145/3627673.3679599
- Yim, T. Y., Tan, W., Zhang, Y., Lam, T. W., & Yiu, S. M. (2025). Demystifying TCFD Disclosures: An AI-Powered Framework for Enhanced Transparency and Trust. *ICAIF 2025 - 6th ACM International Conference on AI in Finance*, 578–586. https://doi.org/10.1145/3768292.3770400
- Zhang, J., Li, Y., Liu, Y. H., & Xie, L. (2024). Research and Practice of NL2SQL Technology Based on LLM for Big Data of Enterprise Finance. *Proceedings - 2024 4th International Conference on Advanced Enterprise Information System, AEIS 2024*, 46–51. https://doi.org/10.1109/AEIS65978.2024.00015