FinIR: The 2nd Workshop on Financial Information Retrieval in the Era of Generative AI

Half-day Workshop, Website: https://finir2025.github.io/

Fengbin Zhu
National University of Singapore
Singapore, Singapore
fengbin@nus.edu.sg

Yunshan Ma
Singapore Management University
Singapore, Singapore
ysma@smu.edu.sg

Fuli Feng
University of Science and Technology of China
Hefei, China
fengfl@ustc.edu.cn

Chao Wang
6Estates Pte Ltd
Singapore, Singapore
wangchao@6estates.com

Huanbo Luan
6Estates Pte Ltd
Singapore, Singapore
luanhuanbo@6estates.com

Guangnan Ye
Fudan University
Shanghai, China
yegn@fudan.edu.cn

Shuo Zhang
Bloomberg
London, United Kingdom
szhang611@bloomberg.net

Dhagash Mehta
BlackRock
New York, United States
dhagash.mehta@blackrock.com

Pingping Chen
Goldman Sachs
Hong Kong, China
pingping.chen@gs.com

Bing Xiang
Goldman Sachs
New York, United States
bing.xiang@gs.com

Tat-Seng Chua
National University of Singapore
Singapore, Singapore
dcscts@nus.edu.sg

Abstract

Recent advancements in Generative AI, such as Large Language Models (LLMs), have demonstrated remarkable success across various general tasks. Extensive studies have explored leveraging generative models in finance, but significant challenges persist. This half-day workshop explores potential approaches and research directions to address these challenges by equipping generative models with advanced Information Retrieval (IR) models. Specifically, this workshop seeks to provide a platform for discussing innovative ideas that facilitate the advancement of IR technology to enrich generative models in finance from four key perspectives: (i) financial IR techniques, (ii) financial IR benchmarking and evaluation, (iii) financial systems and agents/assistants, (iv) and trustworthiness, privacy and security when applying financial IR and generative models. This workshop aims to deepen understanding, accelerate progress, and support the advancement of IR technology to enhance generative models to address financial challenges.

CCS Concepts

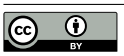
• **Information systems** → **Information retrieval**; • **Computing methodologies** → **Artificial intelligence**.

Keywords

Retrieval Augmented Generation, Finance Information Retrieval, Generative AI, Large Language Models

ACM Reference Format:

Fengbin Zhu, Yunshan Ma, Fuli Feng, Chao Wang, Huanbo Luan, Guangnan Ye, Shuo Zhang, Dhagash Mehta, Pingping Chen, Bing Xiang, and Tat-Seng Chua. 2025. FinIR: The 2nd Workshop on Financial Information Retrieval in the Era of Generative AI: Half-day Workshop, Website: https://finir2025.github.io/ . In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25), July 13–18, 2025, Padua, Italy.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3726302.3730366



SIGIR '25, Padua, Italy

ACM ISBN 979-8-4007-1592-1/2025/07
https://doi.org/10.1145/3726302.3730366

1 Motivation and Scope

Generative models, including large language models [9], multimodal language models [7], and diffusion models [10], have achieved remarkable success in the recent years. Extensive works explore leveraging the power of advanced generative models in completing various tasks in the finance domain, including financial data understanding [14], financial chart analysis [13] and financial synthetic table generation [11], etc. Despite their impressive performance, substantial challenges remains. One primary limitation is that generative models mainly rely on internal parametric knowledge [8], leading to issues like factual inconsistencies and hallucinations, especially when handling knowledge-intensive tasks in finance. Moreover, financial data, derived from diverse sources, is typically heterogeneous and updated at varying frequencies. For instance, updates may occur on daily, seasonal, annual, and random time scales for stock prices, financial research reports, announcements,

economic statistics, and financial news, respectively. Integrating the latest data and knowledge with generative models timely to produce reliable responses is vital.

To tackle these challenges, Retrieval-Augmented Generation (RAG), which integrates Information Retrieval (IR) technology, has emerged as a promising solution and attracted lots of attention in the community. As a premier conference in the field of IR, the SIGIR community has a crucial role in driving advancements in IR research for the finance domain in the era of generative AI. In light of this, this workshop is organized to bring together a diverse group of researchers and practitioners interested in exploring the advancement of IR technologies to enhance generative models in finance. The topics of interest include but are not limited to:

- **Financial IR Techniques**
  - Financial retrieval models including table retrieval, multi-modal retrieval, real-time retrieval, and generative retrieval, etc.
  - Techniques for financial query analysis, query rewriting, query expansion, etc.
  - Techniques for financial data processing, de-duplication, representation, indexing, filtering, and ranking, etc.
- **Financial IR Evaluation**
  - Financial IR tasks and benchmarks for evaluating IR models in finance.
  - Evaluation metrics and methodologies for measuring financial IR models.
  - Evaluation with simulators, such as market simulators.
- **Financial Systems and Agents**
  - Utilizing IR and generative models to develop systems and applications for financial data analysis, stock price prediction, stock recommendation, portfolio optimization, financial content generation, event forecasting, etc.
  - Financial conversational agents and personal financial assistants powered by IR and generative models.
  - Multi-agent systems in finance for algorithmic trading, policy design, behavior modeling, risk assessment, etc.
- **Trustworthiness, Privacy and Security**
  - Exploring and resolving the challenges related to trustworthiness and reliability in applying IR and generative models in finance, including distinguishing AI-generated financial content and ensuring factual consistency, etc.
  - Privacy-preserving access and processing financial data, such as sensitive information anonymization and privacy regulations compliance (e.g., PDPA and GDPR).
  - Financial data security when using IR technology and generative models, such as financial misinformation detection and data leakage prevention.

2 Rationale

2.1 Relevance

This workshop aims to explore the integration of advanced IR technology with generative models to address practical challenges in the finance domain, which aligns seamlessly with the objectives of the SIGIR conference. Moreover, the advancement of IR models over multi-modal and heterogeneous finance data in this workshop has the potential to inform and enhance the incorporation of IR models into generative models across the general domains.

2.2 Objectives and Expected Outcome

This workshop seeks to inspire researchers to explore pioneering avenues to address financial challenges by leveraging IR technology and generative models. These emerging technologies offer the potential to enrich industry applications with distinctive features while encouraging fresh academic research directions. We expect to receive cutting-edge contributions in this promising field with exciting ideas and novel contributions.

2.3 Target Audience

The attractiveness of this workshop stems from its dedication to the evolving areas of FinTech and AI for finance. It aims to engage a diverse audience of researchers and industry professionals at the intersection of AI and finance, providing a distinctive platform for exchanging ideas, methods, and accomplishments. It encourages interdisciplinary collaboration and inspires exploration of novel applications across fields. We anticipate receiving over 20 submissions and having around 40 in-person and 80 online participants.

2.4 Related Workshops

The Workshops Held Before: We initiated our first workshop [6] at SIGIR'20, where we invited esteemed keynote speakers from both academia and industry, including Prof. Longbing Cao from University of Technology Sydney, Dr. Edgar Meij from Bloomberg UK, Dr. Haiqin Yang from Ping An Life Insurance of China and Dr. Ben Tan from WeBank. Due to the COVID-19 pandemic, this workshop was held virtually and successfully attracted over 70 participants via Zoom.

Other Related Workshops: In addition to our past workshop, there are some other related workshops: 1) workshops on "Financial Technology and Natural Language Processing(FinNLP)" [3–5]; 2) workshops on "Economics and Natural Language Processing (ECONLP)"; 3) workshops on "Knowledge Discovery from Unstructured Data in Financial Services (KDF)" [12]; 3) workshops on "Economics and Natural Language Processing (ECONLP)". The FinNLP and ECONLP series workshops emphasize the application of Natural Language Processing (NLP) in the finance industry while the KDF workshops focus on data mining and knowledge discovery from unstructured financial documents. In recent years, the ACM International Conference on AI in Finance (ICAIF) [1, 2] has established itself as an active venue for showcasing research at the intersection of AI and finance. At ICAIF, some related workshops have been held, including "LLMs and Generative AI for Finance" and "Multimodal Financial Foundation Models (MFFMs)" at ICAIF'24. Both workshops place more emphasis on generative and large foundation models. In contrast, our workshop highlights the advances in integrating IR technology with generative models.

2.5 Diversity

Our workshop places a strong emphasis on diversity across multiple dimensions, including the backgrounds of organizing team and invited speakers. The organizing committee comprises members from various institutions spanning East Asia, Southeast Asia, the United Kingdom, and the United States, representing a blend of academic and industry perspectives. We keep gender diversity by including both male and female researchers. Moreover, our organizers bring

Table 1: Program schedule.

Event	Time
Opening Remarks by Co-Chairs	08:30–08:40
Keynote Talk #1 followed by QA	08:40–09:10
Paper Session #1 (2 papers)	09:10–09:40
Keynote Talk #2 followed by QA	09:40–10:10
Paper Session #2 (2 papers)	10:10–10:40
Tea Break	10:40–11:10
Keynote Talk #3 followed by QA	11:10–11:40
Paper Session #3 (2 papers)	11:40–12:10
Closing Remarks by Co-Chairs	12:10–12:20

a wide range of professional expertise, with some specializing in foundational research while others focus on AI applications. This diverse combination of experiences and perspectives enriches the workshop, fostering a collaborative platform that encourages the exchange of a broad spectrum of ideas and expertise.

3 Workshop Program Format

This workshop will be held for **half a day**. We will invite three researchers in the domain of finance and IR, each of whom will give a 30-minute keynote speech. We will also invite several senior researchers and developers to organize a panel discussion on future directions. Besides, in this workshop, we expect to receive a minimum of 10 submissions and accept at least 6 papers. Each paper would be given around 15 minutes for presentation and QA. We present the preliminary program schedule in Table 1. Kindly note that the paper acceptance plan and the program schedule are tentative and may be subject to potential adjustments according to the requirements of the conference chairs.

4 Call for papers

4.1 Introduction

The Workshop on "Financial Information Retrieval in the Era of Generative AI", held in conjunction with the SIGIR 2025 conference, offers a dedicated platform to advance research and innovation in AI for finance. The workshop's goal is to drive IR and generative models toward addressing various real-world challenges in the finance domain. We invite contributions on various aspects of the topic, including future perspectives, technological advancements, practical applications, and more. Join us for insightful discussions and knowledge exchange to shape the future of AI for finance powered by advanced IR technology and generative models.

4.2 Objectives and Scope

This workshop aims to encourage innovative research on advancing financial IR technologies to enrich generative models' capabilities from four key perspectives: (i) financial IR techniques, (ii) financial IR benchmarking and evaluation, (iii) financial systems and agents/assistants, (iv) and trustworthiness, privacy and security when applying financial IR and generative models. We summarize the detailed objectives and scope in Section 1.

4.3 Submission Information

Submission Guidelines: Papers must be submitted as a single PDF file using the ACM SIGIR 2025 template. Submissions may range from 4 to 8 pages, with no limit on the number of pages for references. Authors can choose the appropriate length for their paper, as no distinction is made between long and short submissions. All papers will undergo a "double-blind" review process, following the same evaluation procedure and timeline. Expert peer reviewers will evaluate submissions based on their relevance to the workshop, scientific innovation, and technical quality.

4.4 Participation and Selection Process

Each submission will be assigned to at least three program committee members without conflicts of interest (COI) for review. Once the reviewers' comments are gathered, the organizers will convene a meeting to discuss the reviews and finalize the decision.

4.5 Timeline

The following is the timeline for this workshop.

- Submission deadline: 30 April, 2025
- Paper acceptance notification: 21 May, 2025
- Workshop date: 17 July 2025

5 Organizers

- **Fengbin Zhu**: Dr. Fengbin Zhu is a Research Fellow at the National University of Singapore (NUS). He received his Ph.D. in Computer Science from NUS, supervised by Prof. Tat-Seng Chua. His research interests include Large Language Models (LLMs), Information Retrieval (IR), Document Intelligence (DI), and their applications in finance. He has published several papers in top-tier venues like ACL, AAAI and ACM MM. He also serves as a reviewer and program committee member for conferences and journals such as ACL, AAAI, TKDE, and TOIS.
- **Yunshan Ma**: Dr. Yunshan Ma is a Tenure-track Assistant Professor in school of computing and information systems, Singapore Management University (SMU). He received his PhD degree from the School of Computing, National University of Singapore. His research interests include multimodal event forecasting and recommender systems, especially associated applications in vertical domains, such as macro-level indicator forecasting or micro-level user behavior prediction in finance, economics, and politics, etc. His works have appeared in top-tier conferences and journals such as SIGIR, SIGKDD, WWW, WSDM, TKDE, etc. He received the Best Student Paper Award in ACM ICMR 2021. Moreover, he has served as the reviewers for international conferences and journals including SIGIR, SIGKDD, WWW, TKDE and TMM, etc.
- **Fuli Feng**: Dr. Fuli Feng is a Professor at the University of Science and Technology of China. He earned his Ph.D. in Computer Science from the National University of Singapore in 2019. His research covers information retrieval, data mining, causal inference, and multimedia processing, with over 60 publications in top conferences like SIGIR, WWW, SIGKDD, and journals such as TKDE and TOIS. He received the Best Paper Honorable Mention at SIGIR 2021 and the Best Poster Award at WWW 2018. Dr. Feng has served as a PC member for major conferences/journals, including SIGIR, WWW, SIGKDD, NeurIPS, TKDE, TPAMI, *etc.* He has rich experience in organizing workshops and tutorials at WWW'21&22&24, SIGIR'20&23&24, and RecSys'21.

- **Chao Wang**: Dr. Chao Wang is the CTO of 6Estates Pte Ltd. Previously, he was a senior research scientist at Baidu. He holds a PhD in Computer Science from Tsinghua University. His work has appeared in major journals and conferences such as SIGIR, CIKM, TOIS, and IRJ. His main achievements include 2015 SIGIR Best Paper Honorable Mention Award, 2015 Scientific Technology Advance Award of Beijing City (First Prize), and 2016 CIPS Excellent PHD thesis Award.
- **Huanbo Luan**: Dr. Huanbo Luan is CEO of 6Estates Pte Ltd, a leading Enterprise GenAI company in Singapore. He was the deputy director of NExT++ Research Center established by Tsinghua University and National University of Singapore. He received his Ph.D degree in computer science from Institute of Computing Technology, Chinese Academy of Sciences in 2008. His research interests include artificial intelligence, natural language processing and multimedia information retrieval.
- **Guangnan Ye**: Dr. Guangnan Ye is a Professor from School of Computer Science and Technology, FinTech Institute at Fudan University, as well as FinTech Research Center of the People's Bank of China. He is the chief scientist of the National Key R&D Program. Dr. Ye got his PhD at Columbia University, supervised by Prof. Shih-Fu Chang. His main research areas include FinTech research, machine learning, computer vision, etc.
- **Shuo Zhang**: Dr. Shuo Zhang works as an AI Researcher with Bloomberg's AI Engineering team Group in London. Before then, he obtained his doctoral degree at the University of Stavanger, Norway, in 2019, under the supervision of Prof. Krisztian Balog. He used to be a visiting scholar at Carnegie Mellon University, Pittsburgh, USA, in 2019, visiting Prof. Jamie Callan. His main research interests include information retrieval and text mining.
- **Dhagash Mehta**: Dr. Dhagash Mehta is the Head of Applied Machine Learning Research (Investment Management) at Blackrock Inc. and an Editorial Board Member at the Journal of Financial Data Science and Journal of ESG and Impact Investing (both PMR journals). Previously, he was a Senior Manager, Investment Strategist at Investment Strategy Group at The Vanguard Group, Senior Research Scientist at United Technologies (UTX) Research Center, and a Research Assistant Professor at University of Notre Dame. Dr. Mehta's research areas are machine/deep learning; quantitative finance, and computational mathematics, science and engineering. Dr. Mehta served as General Chair of ICAIF'24, Workshop co-chair at ICAIF'23 and has experience in organizing workshops at NeurIPs, ICAIF, SIAM and AMS.
- **Pingping Chen**: Ms. Pingping Chen is the Head of Applied AI in APAC and Tech Fellow at Goldman Sachs. Before her tenure at Goldman Sachs and leading the Gen AI efforts in APAC, she held key roles in AI/ML at several leading global organizations including Microsoft Research, Yahoo, Baidu, and HSBC. Ms. Chen earned her master's and bachelor's degrees from Tsinghua University in 2011 and 2008, respectively.
- **Bing Xiang**: Dr. Bing Xiang is a seasoned expert in artificial intelligence and machine learning, currently serving as the Head of AI Research and Managing Director at Goldman Sachs, US. Previously, Dr. Xiang led a global team as the Director of Applied Science at AWS AI Labs from 2017 to 2023, where significant AI services such as Amazon Bedrock and Amazon CodeWhisperer were developed. Prior experience includes leadership roles at IBM Watson Group, Thomson Reuters, IBM T. J. Watson Research Center, and BBN Technologies, focusing on cognitive analytics, deep learning, and natural language processing. Dr. Xiang holds a Ph.D. in Electrical and Computer Engineering from Cornell University and two degrees from Peking University in Signal and Information Processing and Electronics and Information Systems.
- **Tat-Seng Chua**: Dr. Tat-Seng Chua is the KITHCT Chair Professor at the School of Computing, National University of Singapore (NUS). Dr. Chua was the Founding Dean of the School of Computing from 1998-2000. His main research interests include unstructured data analytics, video analytics, conversational search and recommendation, and robust and trustable AI. Dr. Chua is the recipient of the 2015 ACM SIGMM Achievements Award, and the winner of the 2022 NUS Research Recognition Award.

References

[1] 2023. *ICAIF '23: Proceedings of the Fourth ACM International Conference on AI in Finance* (Brooklyn, NY, USA). Association for Computing Machinery, New York, NY, USA.

[2] 2024. *ICAIF '24: Proceedings of the 5th ACM International Conference on AI in Finance* (Brooklyn, NY, USA). Association for Computing Machinery, New York, NY, USA.

[3] Chung-Chi Chen, Hen-Hsen Huang, Hiroya Takamura, and Hsin-Hsi Chen (Eds.). 2022. *Proceedings of the Fourth Workshop on Financial Technology and Natural Language Processing (FinNLP).* Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (Hybrid).

[4] Chung-Chi Chen, Tatsuya Ishigaki, Hiroya Takamura, Akihiko Murai, Suzuko Nishino, Hen-Hsen Huang, and Hsin-Hsi Chen (Eds.). 2024. *Proceedings of the Eighth Financial Technology and Natural Language Processing and the 1st Agent AI for Scenario Planning.* -, Jeju, South Korea.

[5] Chung-Chi Chen, Hiroya Takamura, Puneet Mathur, Remit Sawhney, Hen-Hsen Huang, and Hsin-Hsi Chen (Eds.). 2023. *Proceedings of the Fifth Workshop on Financial Technology and Natural Language Processing and the Second Multimodal AI For Financial Forecasting.* -, Macao.

[6] Fuli Feng, Cheng Luo, Xiangnan He, Yiqun Liu, and Tat-Seng Chua. 2020. FinIR 2020: The First Workshop on Information Retrieval in Finance. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, China) *(SIGIR '20).* Association for Computing Machinery, 2451–2454.

[7] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. 2024. Visual instruction tuning. *Advances in neural information processing systems* 36 (2024).

[8] Alex Mallen, Akari Asai, Victor Zhong, Rajarshi Das, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. When Not to Trust Language Models: Investigating Effectiveness of Parametric and Non-Parametric Memories. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).* Association for Computational Linguistics, 9802–9822.

[9] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 35 (2022), 27730–27744.

[10] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 10684–10695.

[11] Timur Sattarov, Marco Schreyer, and Damian Borth. 2023. FinDiff: Diffusion Models for Financial Tabular Data Generation. In *Proceedings of the Fourth ACM International Conference on AI in Finance* (Brooklyn, NY, USA) *(ICAIF '23).* Association for Computing Machinery, New York, NY, USA, 64–72.

[12] Sameena Shah, Xiaodan Zhu, Wenhu Chen, Manling Li, Armineh Nourbakhsh, Xiaomo Liu, Zhiqiang Ma, Charese Smiley, Yulong Pei, and Akshat Gupta. 2023. Knowledge Discovery from Unstructured Data in Financial Services (KDF) Workshop. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan) *(SIGIR '23).* Association for Computing Machinery, New York, NY, USA, 3464–3467.

[13] Ziao Wang, Yuhang Li, Junda Wu, Jaehyeon Soon, and Xiaofeng Zhang. 2023. FinVis-GPT: A Multimodal Large Language Model for Financial Chart Analysis. arXiv:2308.01430 [cs.CL]

[14] Fengbin Zhu, Ziyang Liu, Fuli Feng, Chao Wang, Moxin Li, and Tat Seng Chua. 2024. TAT-LLM: A Specialized Language Model for Discrete Reasoning over Financial Tabular and Textual Data *(ICAIF '24).* Association for Computing Machinery, 310–318.