

FinNLP 2026

**The 11th Workshop on Financial Technology and Natural
Language Processing**

Proceedings of the Workshop

October 28, 2026

The FinNLP organizers gratefully acknowledge the support from the following sponsors.

Bloomberg
Engineering



©2026 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-466-8

Introduction

Welcome to the 11th Workshop on Financial Technology and Natural Language Processing (FinNLP 2026), held in Budapest, Hungary, on October 28, 2026, co-located with the 2026 Conference on Empirical Methods in Natural Language Processing (EMNLP 2026).

FinNLP brings together researchers and practitioners working at the intersection of natural language processing and finance. The workshop invites work on artificial intelligence, machine learning, natural language processing and understanding, big data, statistical learning, data analytics, and deep learning, with a focus on knowledge discovery in the financial services domain. Topics of interest this year included language modeling on financial corpora, including tabular, numerical, and multi-modal data; graph representation learning and mining on financial data; financial search and question answering systems; event discovery and its impact on equity prices; large language model applications in financial auditing, compliance, and regulatory reporting; hallucination mitigation and evaluation; interpretability and explainability of large language models; and the ethical implications and bias mitigation of AI applications in finance.

This year we accepted two types of submissions: direct submissions to the workshop and papers committed through ACL Rolling Review (ARR). We received a total of 172 submissions (165 direct submissions and 7 ARR commitments), of which 151 entered the review process after desk rejections and withdrawals: 96 long papers, 31 short papers, 4 demo papers, 6 ARR commitments, and 14 system description papers for the FinEval 2026 shared task. Each direct submission typically received three reviews from the program committee, with area chairs overseeing the discussion and making recommendations; ARR commitments were evaluated on the basis of their ARR reviews and meta-reviews. After careful consideration, we accepted 62 research-track papers (44 long, 13 short, 1 demo, and 4 ARR commitments), for an acceptance rate of 45%, together with 14 shared task system description papers, for a total of 76 accepted papers. Of these, 12 were selected for oral presentation and 64 for poster presentation. At the request of their authors, 9 accepted papers are presented at the workshop as non-archival papers and are not included in this volume, and one accepted paper was withdrawn. These proceedings also include the overview paper of the FinEval 2026 shared task written by its organizers, and therefore contain 67 papers.

In addition to the research track, FinNLP 2026 hosted the FinEval 2026 shared task on Multilingual Financial Language Understanding, with three subtasks: Greek financial named entity recognition over company filings, Japanese implicit communication recognition in company Q&A, and cross-document multilingual financial question answering combining SEC filings with news. We thank the shared task organizers and all participating teams; the shared task overview and the system description papers are included in these proceedings.

We are honored to have Daniel Preoțiuc-Pietro (Bloomberg) and Bhaskarjit Sarmah (Domyne) as our keynote speakers, and Yohei Seki (University of Tsukuba), Prayag Tiwari (Halmstad University), Jacob Chanyeol Choi (LinqAlpha), and Qianqian Xie (Wuhan University) as invited speakers. The program further includes two oral sessions, a poster session, the FinEval 2026 shared task session with the award presentation, and a panel discussion.

We would like to thank all the authors for submitting their work, the members of the program committee for their thorough and timely reviews, and the organizing committee for their efforts in making this workshop possible. We are also grateful to our sponsors, Bloomberg and Domyne, for their generous support, including the Domyne Best Paper Award, and to the EMNLP 2026 workshop chairs and the ACL publication chairs for their guidance.

We hope you enjoy the workshop and find the papers in these proceedings inspiring.

Jimin Huang, Yongjae Lee, Sophia Ananiadou, Hsin-Hsi Chen, Hiroya Takamura, and Chung-Chi Chen
General Chairs, FinNLP 2026

Organizing Committee

General Chairs

Jimin Huang, University of Manchester, UK and The FinAI, USA
Yongjae Lee, UNIST, South Korea
Sophia Ananiadou, University of Manchester, UK
Hsin-Hsi Chen, National Taiwan University, Taiwan
Hiroya Takamura, AIST, Japan
Chung-Chi Chen, National Institute of Informatics (NII), Japan

Publicity Chair

Xueqing Peng, Yale University, USA

Shared Task Chair

Zhiwei Liu, University of Manchester, UK

Sponsorship Chair

Yankai Chen, MBZUAI, UAE

Review Chair

Lingfei Qian, Yale University, USA

Publication Chair

Yan Wang, Yale University, USA

Local Chair

Zhuohan Xie, MBZUAI, UAE

Program Committee

Area Chairs

Nino Antulov-Fantulin, ETH Zurich and Aisot Technologies
Yupeng Cao, Yale University
Yankai Chen, Mohamed bin Zayed University of Artificial Intelligence
Xutong Liu, University of Twente
Zhiwei Liu, University of Manchester
Xueqing Peng, Yale University
Lingfei Qian, Yale University
Prayag Tiwari, Halmstad University
Yan Wang, Yale University
Zhuohan Xie, Mohamed bin Zayed University of Artificial Intelligence
Yong Xie, Amazon
Jiangpeng Yan, E Fund Management Co., Ltd.
Shiqi Ye, Academy of Mathematics and Systems Science, Chinese Academy of Sciences

Reviewers

Ahmed Abdullah, Piyush Agarwal, Zohair Ahmed, Jubayer Al Mahmud, Areej AlOtaibi, Seyed Mosayeb Alam, Paulo Alencar, Avinash Goutham Aluguvelly, Tasniem Nasser Alyahya, Diego Amaya, A S Aravindhakshan, Elliott Ash, Atousa Assadihaghi, Alankar Atreya

Max Baak, Shubham Barge, Devesh Batra, Ofir Ben Shoham, Morris Groot Beumer, Abhishek Bhandari, Zacchaeus Bolaji

Xin Cai, Dejun Cao, Mingyu Cao, Wenbo Cao, Yupeng Cao, Emma Ceccherini, Nischal Reddy Chandra, Ben Charoenwong, Bowei Chen, Pin-Hao Chen, Eunjung Cho, Wooah Choi, Saebyeok Chu, Arnau Rodon Comas, Vignesh Coumarane, Ziqiang Cui, Jannic Alexander Cutura

Giulia D'Agostino, Sridhar Dasaratha, Yu-Chen Den, Suryadeep Singh Deswal, Durjoy Dey, Sasawat Dhanuka, Fabrizio Dimino, Valeria Vieira dos Santos

Liancheng Fang, Fatema Tuj Johora Faria, Dongji Feng, Weibin Feng

Stoyan Ganchev, Somraj Gautam, Daria Godorozha, Jie Gong, Runteng Guo, Wei Guo, Nyida Gyal

Mehmet Haklidi, Krishna Harish, Chetan Harsha, Wanying He, Yueru He, Yuxian He, Amr Hegazy, Simon Heininger, Shu-Kai Hsieh, Jingying Hu, Caishuang Huang, Ke-Wei Huang, Xianting Huang, Xiyang Huang, Yu-Shiang Huang

Apoorva Jaiswal, JoonHowi Jeon, Denglin Jiang, Yuechen Jiang

Anup K. Kalia, Anusha Kamath, Hong Kang, Munyoung Kang, Alireza Kargarzadeh, Arman Khaledian, Nariman Khaledian, Genady Kogan, Ziyang Kuang, Shamik Kundu, Kseniia Kuvshinova

Victoire Laurent, Liliya Lechner, Yesol Lee, Yunyoung Lee, Baichuan Li, Haohang Li, Mingqi

Li, Peng Li, Yifan Li, Zhentao Li, Fangfang Lin, Ping-Sheng Lin, Yuzhu Ling, Haifeng Liu, Yontai Liu, Zhiwei Liu, Zhuosheng Liu, Giuliano Lorenzoni, Peng Lu, Fabian Lukassen

Jintao Ma, Linhai Ma, Peixian Ma, Prabir Mallick, Anar Mehraliyev, Vaibhav Mehrotra, Ge Meng, Yanlong Meng, Mukaffi Bin Moin

Ann Naser Nabil, Vu Gia Nam, Aravind Narayanan, Chun Chet Ng, Jian-Yun Nie, Talapaka Nithin, Erfan Nourbakhsh, Valentin Noël

Rapheal Obodugo, Mateusz Ostaszewski

Rajveer Singh Pall, Nisarg Patel, Ying Peng, Adam Poliak, Shivika Prasanna

Wenjie Qiu, Yunzhi Qiu

Aulia Kharis Rakhmasari, Shashishekar Ramakrishna, Aman Rangapur, Suprabhat Rapolu

Arush Sachdeva, Yunus Santur, Krish Sapru, Anirban Sen, Théo Nicolas Sitjar, Kyungsub Song, Yanlin Song, Laven Srivastava, Varad Srivastava, Kang Sumin, Yan Sun, Sajag Swami

Yu-Chien Tang, Aaditya Anand Thokal, Lara Turgut, Oddharak Tyagi

Hafeez Ullah, Takehito Utsuro, Nong Nhu Uyen

Natalia Vanetik, Ishan Verma, Shivansh Verma, Natan Vidra

Qianhui Stella Wan, Chuan-Ju Wang, Dong Wang, Wilson Wang, Yinzhi Wang, Yuyan Wang, Zhuo Wang, Zining Wang, Wei Wei, Xinyu Wei, Lars Wilhelmi, Haolun Wu, Wenbin Wu, Yue Wu, Zongxiao Wu

Xinyi Xia, Ruoyu Xiang, Anke Xu, Shijie Xu, Xinhe Xu, Yihuai Xu

Lalit Yadav, Yuhong Yan, Junyi Yao, Ping Yao, Xuan Yao, An-Zi Yen, Simge Yildirim, Shixiang You, Dianshi Yu, Honghai Yu, Han Yuan

Huaye Zeng, Bo Zhang, Enze Zhang, Fan Zhang, Nathan Zhang, Vincent Jim Zhang, Xilun Zhang, Xinni Zhang, Yunbai Zhang, Nan Zhao, Xinyuan Zhao, Yingzhu Zhao, Jin Zheng, Yixiang Zheng, Zihao Zheng, Yixi Zhou, Liyun Zhu, Mingxi Zou

Table of Contents

<i>Definitional or Computational? Decomposing LLM Numeric Fidelity in Financial-Statement Analysis with an XBRL Cascade Oracle</i>	
Oussama Chahed	1
<i>Scripted Hesitation in a Financial Customer-Service Chatbot: A Brazilian Portuguese Case Study</i>	
Valeria Vieira dos Santos	10
<i>FinLCR: Scaling Financial Long-Context Reasoning via Curriculum Reinforcement Learning</i>	
Wenjie Qiu, Yong Xie, Karan Aggarwal, Aitzaz Ahmad and Stephen Lau	15
<i>Can LLMs Catch a Rigged Backtest? A Clean-Control Calibration Benchmark</i>	
Makar Ulesov, Vladislav Smirnov, Omar Ibrahim and Arsenii Bobovnikov	26
<i>FinEvo: Financial Language Agents Shape the Markets That Evaluate Them</i>	
Mingxi Zou, Jiaxiang Chen, Langzhang Liang, Zhuo Wang, Aotian Luo, Chi Zhang, Danny Dongning Sun and Zenglin Xu	31
<i>TRUST-BENCH: A Synthetic Japanese Financial-Group Benchmark for Enterprise Document Access in Coding Agents</i>	
Shamik Kundu, Anand Mekala and Kazutaka Yamasaki	53
<i>One Advisor, Many Faces: How User Personas Shape LLM Stock Recommendations</i>	
Yuxian He	65
<i>Diagnosing Fine-Grained Inconsistency Classification in Financial Disclosure Text</i>	
Aman Kumar, Lasitha Vidyaratne, Dipanjan Dipak Ghosh, Arnab Chakrabarti and Ahmed K. Farahat	75
<i>Trust the Query, Not the Answer: Auditable Text-to-SPARQL over the GLEIF Register</i>	
Jannic Alexander Cutura	91
<i>MortarBench: Evaluating Mortgage Loan Origination Agents</i>	
Matthew Toles, Yunan Lu, Manav Munjal, Bojun Liu, Yuanhao Deng, Stephanie Selig, Derek Rindner, Cheng Li and Zhou Yu	98
<i>CARPAS: Towards Content-Aware Refinement of Provided Aspects for Summarization in Large Language Models</i>	
Yong-En Tian, Yu-Chien Tang, An-Zi Yen and Wen-Chih Peng	110
<i>What the Window Does Not Contain: Auditing Provenance in a Document-Grounded Instability Benchmark</i>	
Seyed Mosayeb Alam	125
<i>Auditing Bank Statement PDFs: A Multi-Tier Parser and Design Choices for the NL-to-SQL Query Layer Above It</i>	
Anamul Hoque Emtiaj and Anindya Iqbal	142
<i>SpendSight: Privacy-Preserving Personal Finance Q&A with Sub-2B LLMs</i>	
Sajag Swami and Yash Pandey	153
<i>CORAL: Closed-loop Oracle Alignment for Agentic Tax Engine Modernization</i>	
Maarten De Raedt, Ahmed Abdullah and Abhinav Agrawal	180

<i>What Deterministic Verification Actually Fixes in Financial QA: A Diagnostic Analysis of Rule Heterogeneity Across Datasets</i>	
Aaditya Anand Thokal	193
<i>Where A Small Language Model Helps in Invoice Categorisation, Understood Through Embedding Geometry</i>	
Emma Ceccherini, Daniel John Lawson and Anjulika Salhan	201
<i>FinDRA: Multi-Agent Financial Reasoning under Asymmetric Information and Unequal Authority</i>	
Inyoung Won and Yunyoung Lee	224
<i>FRED-SemBench: Evaluating Semantic Reliability in LLM Access to Macroeconomic Data</i>	
Wilson Wang, Chandler Han and Peter Zhang	237
<i>BankShield: Fairness-Aware Safety Guardrails for Indian-Language Banking AI</i>	
Piyush Agarwal	243
<i>NLP for Crowdfunding Success Prediction: A Systematic Mapping Review Exposes an Underexplored Equity Domain</i>	
Chaya Liebeskind, Hana Tzur and Shmuel Liebeskind	263
<i>When Words Move Markets: Sentiment and Semantic Complexity in Central Bank Communication and Equity Volatility</i>	
Arush Sachdeva and Anirban Dasgupta	274
<i>Same Financials, Different Identity: A Counterfactual Audit of Demographic Sensitivity in LLM-Based Credit-Risk Scoring</i>	
Mustafa Furkan Zenginoğlu, Zeynep Beyda Kara, Hasan Fatih Durkaya, Mehmet Haklıdır and Tuna Çakar	313
<i>CAST: An Interpretable Evidence-Based Confidence Estimation Tool for RAG with Human-in-the-Loop Verification</i>	
Xutong Liu, Morris Groot Beumer, Joerg Robert Osterrieder and Max Baak	332
<i>GRPO for Financial Advice Generation: Outperforming Commercial LLMs under CATE Evaluation</i>	
Ofir Ben Shoham, Shrutendra Harsola, Vignesh Subrahmaniam, Shravan Mohan, Yakov Gazman and Oded Vainas	342
<i>LatentDemandBench: A Contamination-Resistant Diagnostic for Pre-Recognition Product–Need Ranking in Synthetic Economies</i>	
Avinash Goutham Aluguvelly	351
<i>What Ensemble Brier Scores Hide: Evidence Routing at a Fixed Article-Exposure Budget in LLM Prediction-Market Forecasting</i>	
Avinash Goutham Aluguvelly	373
<i>Continuous Financial Knowledge Graph Construction with Citation-Grounded Agentic Question Answering</i>	
Arnau Rodon Comas	383
<i>Relevant Is Not Enough: Expert-Audited Decision-Grade Web Evidence for Financial Fact Verification</i>	
Hyunkyung Park, Eui Hyun Kim, Wooah Choi and Arkaitz Zubiaga	394
<i>Helping Customers in Distress: An LLM-powered Agent that Converses, Probes, and Routes</i>	
Alankar Atreya, Stefan Sylvius Wagner, Devesh Batra, Robert Hankache, Cristovão Iglesias Jr, Patrick Sinclair, Giulio Pelosio, Michael McMillan, Greig A Cowan and Raad Khraishi	403

<i>Beyond Compressibility: Transferable Low-Dimensional Structure and the Limits of Compression in Financial Retrieval</i>	
Oddharak Tyagi, Shashishekar Ramakrishna, Chetan Harsha and Sridhar Dasaratha	408
<i>Same Meaning, Different Wording: Evaluating Paraphrase Robustness in Financial Natural Language Inference</i>	
Yike Li and Adam Poliak	424
<i>Regression-Aware Statistically Gated Policy Updating for Korean Financial LLM Input Guardrails</i>	
Eunjung Cho, Dae-Young Park and Minwoo Kim	434
<i>Characterizing Ambiguity in Financial Document Question Answering</i>	
Kelsey Rook, Clare Arrington, Tomek Strzalkowski and Razvan Bunescu	441
<i>How LLM Financial Agents Respond to Stated Affect Depends on Their Assigned Role</i>	
Da Chen Lian, Yu-Lin Chang, Yen-Hsi Huang and Shu-Kai Hsieh	454
<i>Preliminary Study on LLM-as-a-Judge Evaluation of Causal Mechanism Explanations in Financial News and its Bias Analysis</i>	
Yucheng Xie, Yurina Kosai, Rikuto Tsuchida and Takehito Utsuro	469
<i>Valid but Wrong: Domain Invariants and a Cross-Model Plan-Validity Benchmark for Household-Finance LLMs</i>	
Vignesh Coumarane	486
<i>CLAIR-Fin: An Adversarial Multi-Agent Framework for Claim-Level Verification and Adaptive Debate in Cross-Modal Financial QA</i>	
Fatema Tuj Johora Faria, Mukaffi Bin Moin, Jubayer Al Mahmud, M. F. Mridha and Md. Alam Hossain	491
<i>When FEVER Meets Finance: A Controlled Benchmark Reveals Overlap-Driven Financial Claim Verification</i>	
Suprabhat Rapolu	518
<i>Perceived Professionalism in Financial Inquiry: An Interpretable and Empirically Grounded Linguistic Modeling of Textual Preference</i>	
Giulia D’Agostino and Chung-Chi Chen	524
<i>Selective Compliance in Financial Rule Reasoning: A Solver-Verified Chinese Benchmark and Empirical Analysis</i>	
Peng Li and Yinchun Wang	549
<i>Hidden Below the Landing Page: Tracking Linked Pages and Relation-Centered DOM Summaries for Financial Phishing Detection</i>	
Hyeji Heo, Dae-Young Park and Minwoo Kim	566
<i>Prediction-Market Seed Capital Recovery from Noise-Dominant Flow</i>	
Yankai Chen, Bowei He, Zhuohan Xie, Zhiwei Liu, Fan Zhang, Xueqing Peng, Yan Wang, Lingfei Qian and Xue Liu	574
<i>DeXposure-Claw: An Agent System for DeFi Risk Supervision</i>	
Aijie Shu, Bowei Chen, Wenbin Wu, Cathy Yi Hsuan Chen and Fengxiang He	590
<i>DelistBench: Evaluating Search-Enabled LLMs for Auditable Corporate-Event Database Completion</i>	
Xuan Yao, Shuping Li, Yang Dai, Yi Zhou and Ke-Wei Huang	601

<i>DealTrace: Localizing and Intervening on Failures in Agentic Financial Reasoning</i>	
Alina Hyk	611
<i>Topic-Aware Environmental Claim Extraction from Corporate News</i>	
Tanay Kumar Gupta, Vaishali Aggarwal and Ishan Verma	632
<i>Do Tools Matter, or Just Their Descriptions?</i>	
Ananya Tomar and Ayush Nangia	650
<i>Human Agreement and Return Association Are Not Interchangeable Criteria</i>	
A S Aravinthakshan, Laven Srivastava and Harsh Nandwani	667
<i>Bags of Words, Bags of Analysts: Two Event Spaces for Industry Classification</i>	
Artem Lensky	681
<i>Market Signal Injection: Adversarial Context Manipulation of LLM Pricing Agents</i>	
Dohun Lee and Hyunwoo Park	687
<i>EVOQUANT: Self-Evolving Verifier-Guided Strategy Optimization for Robust Quantitative Trading</i>	
Jie Mao, Changlun Li, Xiang Li, Qiqi Duan, Jinhui Yuan, Xiang Liu, Yuyu Luo, Jing Tang and Nan Tang	717
<i>FinEval Challenge @ EMNLP 2026 - FinNLP Workshop: Multilingual Financial Language Understanding</i>	
Yupeng Cao, Haohang Li, Yuechen Jiang, Yueru He, Yuyan Wang, Ruoyu Xiang and Yangyang Yu	730
<i>PIO at the FinEval 2026: An OpenRouter API Pipeline for Multilingual Financial Understanding</i>	
Israel Olanrewaju Odeajo	740
<i>Maddies NLP at the FinEval 2026: Proposition-Aware Tagged Reasoning for Japanese Implicit Communication Recognition</i>	
Madhav Kapoor, Aditya Channa and Jastej Singh	745
<i>ADML at the FinEval 2026: Version-Aware Retrieval and Grounded Sparse Editing for Multilingual Financial QA</i>	
Gautier Marti and Khaled AlNuaimi	750
<i>AhsanBLAH at the FinEval 2026: Grounded Two-Stage Financial QA with LLM-Judge Verification</i>	
Muhammad Ahsan Atiq and Mohib Ahmad Butt	755
<i>Finaira at the FinEval 2026: Dense Retrieval and Validated Prompt Engineering for Cross-Lingual Financial Question Answering</i>	
Yara Mohamed Nasser and Ashraf Mohamed	762
<i>shamikkundu at the FinEval 2026: Zero-Parameter Confidence Elicitation for Japanese Implicit Communication Recognition</i>	
Shamik Kundu	770
<i>Mircea Andrei Razvan at the FinEval 2026: A Unified Prompting Approach for Multilingual Financial Q&A</i>	
Mircea Andrei Razvan	778
<i>Daikoku at the FinEval 2026: Multi-Stage Labelling with Forced-Choice Adjudication</i>	
Kyosuke Morita	783
<i>LangKG at the FinEval 2026: Answer-Contract Knowledge over Instruction Engineering</i>	
Shivika Prasanna, Jaeseong Lee, Hui Su and Yourong Xu	793

<i>Jetsons at the FinEval 2026: Minimal-Pair Augmentation and Label-Routed Ensembling for Greek and Japanese Financial NLP</i>	
Parag Pravin Dakle, Sai Krishna Rallabandi and Preethi Raghavan	800
<i>Fin_Agent at FinEval 2026: Task-Aware Methods for Multilingual Financial NLP</i>	
Mohotarema Rashid, Tanzina Hossain, Xiaoying Song and Lingzi Hong	811
<i>Token Crit at the FinEval 2026: Role-Specialized, Validation-Gated LLM Workflows for JF-ICR and Multilingual Financial Q&A</i>	
Anke Xu and Wenbo Cao	822
<i>Al-Qahera at the FinEval 2026: GEPA Prompt Evolution for Greek Financial NER and Japanese Implicit Communication Recognition</i>	
Mohamed Mohamed and Hazem Ajlan	829
<i>HX-Fin at FinEval 2026: Precision-Anchored Answer Refinement and Multi-Model Evidence Selection for Multilingual Financial Question Answering</i>	
Peng Li and Yinchun Wang	835