findings 2026

The 15th International Joint Conference on Natural Language Processing and the 5th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics

Findings 2026

November 6-10, 2026

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

Co-organizer



Platinum



Gold





Venue Supporter



©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 979-8-89176-444-6

Message from the General Chair

Welcome to AACL-IJCNLP 2026 in Hengqin, China!

This joint conference brings together the Asia-Pacific and international natural language processing communities to share research, exchange ideas, and build connections. We look forward to conversations that connect languages, disciplines, institutions, and different perspectives on the future of our field.

We hope that the conference will provide a welcoming setting for presenting your work, learning from colleagues, and meeting new collaborators. Whether you are joining us in Hengqin or participating online, thank you for contributing to this community.

We thank the organizing committee, reviewers, tutorial instructors, workshop organizers, invited speakers, volunteers, sponsors, and all participants for the work that makes this meeting possible. We wish you a rewarding conference and an enjoyable stay in Hengqin.

Yue Zhang, General Chair
Ruifeng Xu, Ayu Purwarianti, Tong Rong, and Jun Suzuki, Program Chairs

Message from the Local Organizing Chairs

We are delighted to welcome you to the 5th Asia-Pacific Chapter of the Association for Computational Linguistics and the 15th International Joint Conference on Natural Language Processing. It is our great pleasure to host the AACL-IJCNLP community in Hengqin, a dynamic area of the Greater Bay Area just across from Macau.

We sincerely thank the General Chair, Program Chairs, organizing committee members, reviewers, volunteers, sponsors, and all contributors for their dedication and support in making this conference possible. We are also grateful to all authors, speakers, and participants for bringing their work, ideas, and energy to AACL-IJCNLP 2026.

We hope your time at AACL-IJCNLP 2026 will be both productive and enjoyable. Beyond the technical program, we encourage you to experience the local culture and hospitality of Hengqin and the surrounding region, including nearby Macau. We wish you inspiring discussions, fruitful collaborations, and a memorable conference experience.

AACL-IJCNLP 2026 Local Organization Chairs

Derek F. Wong (University of Macau)
Haofen Wang (Tongji University)

Organizing Committee

General Chair

Yue Zhang, Westlake University, China

Program Chairs

Ruifeng Xu, Harbin Institute of Technology, Shenzhen, China
Ayu Purwarianti, Institut Teknologi Bandung, Indonesia
Tong Rong, Singapore Institute of Technology, Singapore
Jun Suzuki, Tohoku University, Japan

Local Organizing Chair

Derek F. Wong, University of Macau
Haofen Wang, Tongji University, China

Local Organizing Committees

Runzhe Zhang, University of Macau
Zhiwen Xie, Central China Normal University, China

Workshop Chair

Xuefeng Bai, Harbin Institute of Technology, Shenzhen, China
Ratish Puduppully, IT University of Copenhagen, Denmark

Tutorial Chair

Zhongqing Wang, Soochow University, China
Naoaki Okazaki, Institute of Science Tokyo, Japan

Demonstration Chair

Ningyu Zhang, Zhejiang University, China
Derry Wijaya, Monash University, Indonesia

Student Research Workshop Chair

Yuxin Cao, Fudan University, China
Shio Hong, Fudan University, China
Yuhang Zhou, Shanghai Innovation Institute, China
Wenhe SUn, Fudan University, China

Best Paper Award Chair

Rui Yan, Wuhan University, China

Publication Chair

Junxian He, Hong Kong University of Science and Technology
Thuy-Trang Vu, Monash University, Australia

Handbook Chair

Leiyang Cui, Tencent AI Lab, China

Publicity Chair

Aditya Joshi, The University of New South Wales, Australia
KyungTae Lim, Korea Advanced Institute of Science Technology, South Korea

Web and App Chair

Xinchi Chen, Fudan University, China
Piji Li, Nanjing University of Aeronautics and Astronautics, China
Shudong Liu, University of Macau

Finance Chair

Hou Pong Chan, Alibaba DAMO Academy

Registration Chair

Wenxuan Zhang, Singapore University of Technology and Design, Singapore

Ethics Chair

Yangqiu Song, Hong Kong University of Science and Technology
Ekaterina Vylomova, The University of Melbourne, Australia

Diversity Inclusion Chair

Mamoru Komachi, Hitotsubashi University, Japan
Emily Ohman, Waseda University, Japan

Student Volunteer Chair

Juntao Li, Soochow University, Japan

Sponsorship Chair

Yan Song, University of Science and Technology of China, China
Yuanhe Tian, Zhongguancun Institute of Artificial Intelligence, China

Keynote Talk

NLP for Mental Health: Risk, Promise, Responsibility

Iryna Gurevych
Technical University of Darmstadt
2026-11-06 09:30:00 – Room: **Muhe B+C**

Abstract: More than a billion people live with a mental health condition, yet care remains scarce and often a poor fit for the person receiving it — so people increasingly turn to unsupervised chatbots to fill the gap. Because language is the richest signal we have into someone's psychological state, natural language processing is the key to doing this better. In this talk, I argue that NLP can genuinely help — but solving it takes clinical grounding, real data, and responsible evaluation, not just larger models. Drawing on work from my group, I show how large language models can turn therapy transcripts into treatment-relevant structure, how clinically grounded synthetic data eases the field's privacy bottleneck, and why evaluation must reflect what clinicians actually care about — not just what's easy to benchmark. I close with an invitation to build mental health AI that is personalized, privacy-aware, and responsibly evaluated: tools that support clinicians, not replace them.

Bio: Iryna Gurevych is a full professor of computer science at the Technical University of Darmstadt, Germany, where she founded and heads the Ubiquitous Knowledge Processing (UKP) Lab. She is a pioneer of computational argumentation and cross-document text modeling, with landmark applications to fact verification and misinformation detection, and her research extends to large language models and interdisciplinary NLP for the social sciences and humanities. Prof. Gurevych served as President of the Association for Computational Linguistics (ACL) in 2023 and is an ACL Fellow, a member of the German National Academy of Sciences Leopoldina and Academia Europaea, and a recipient of the 2025 Royal Society Milner Award.

Keynote Talk

Beyond Self-Play: Towards Effective Self-Evolution in LLMs

Yulan He
King's College London
2026-11-06 14:00:00 – Room: **Muhe B+C**

Abstract: Can large language models keep improving through experience they generate themselves? Self-play and self-generated experience offer promising routes. But producing more synthetic data does not necessarily lead to sustained model capability growth. This talk argues that effective self-evolution depends on whether the learning pipeline continues to provide learnable information. Drawing on our recent work, I will examine several ways such information can arise, including generating informative self-play experience, learning directly from environment-generated experience, and actively seeking novel evidence rather than waiting for tasks to be provided. These studies suggest that effective self-improvement requires more than generating its own experience; it needs experience that contains useful information that the model can actually learn from.

Bio: Yulan He is a Professor in Natural Language Processing at King's College London and a Turing AI Fellow. Her research focuses on making Large Language Models (LLMs) more reliable, interpretable, and safe, particularly in reasoning and agentic AI, with applications spanning across education, health, and science. She has received several awards for her research. Her five-year Turing AI Fellowship project was awarded the Best Research Project (Research Excellence) at the RAi UK AI Robotics Awards 2026. She also received a Best Research Paper Award at the same awards ceremony. In addition, she is the recipient of a SWSA Ten-Year Award and a CIKM Test-of-Time Award.

Keynote Talk

Uncertainty, Belief, and Trust in Language Models

Tim Baldwin
MBZUAI
2026-11-07 09:00:00 – Room: **Muhe B+C**

Abstract: Language models and other AI technologies are being used in an ever-increasing range of applications, including being trialled for decision support or automated decision-making in sensitive applications. A critical component of any such usage is the capability for the model to output a confidence/uncertainty assessment associated with any output, or (statistical) belief underlying a particular action. In this talk, I will present recent work on uncertainty estimation and belief updates, with a particular focus on practical utility in improving model performance and efficiency, and enabling humans to perform tasks more effectively.

Bio: Tim Baldwin is Professor of Natural Language Processing at Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), in addition to being a Melbourne Laureate Professor in the School of Computing and Information Systems at The University of Melbourne, and Co-founder and Chief Scientist of startups including LibrAI (focused on AI safety) and Examn AI (focused on dialogue-based candidate profiling). At MBZUAI he has also held positions including Provost, Associate Provost, and Foundation Chair of the Department of Natural Language Processing.
Tim completed a BSc(CS/Maths) and BA(Linguistics/Japanese) at The University of Melbourne in 1995, and an MEng(CS) and PhD(CS) at the Tokyo Institute of Technology in 1998 and 2001, respectively. He joined MBZUAI at the start of 2022, prior to which he was based at The University of Melbourne for 17 years. His research has been funded by organisations including the Australian Research Council, Google, Microsoft, Amazon, Xerox, ByteDance, SEEK, NTT, and Fujitsu. He is the author of over 500 peer-reviewed publications across diverse topics in natural language processing and machine learning, and the recipient of a number of awards at top conferences.

Keynote Talk

The Genie of the LLM: What Shall We Wish for the Future of NLP?

Nizar Habash
New York University Abu Dhabi
2026-11-08 09:00:00 – Room: **Muhe B+C**

Abstract: Large language models have given NLP extraordinary new capabilities. The genie is out of the lamp. The question is no longer only what it can do, but what we should ask it to do, how we know we got what we really wanted, and what it will do to us, our research, and our education.
I reflect on these questions through a multilingual journey from early work on English, Spanish, and Chinese to decades of Arabic NLP. Arabic provides a revealing case study: its morphology, dialects, scripts, and linguistic variation expose assumptions that broad multilingual experiments can overlook. Recent work illustrates both the remarkable capabilities of LLMs and the continuing value of other models, linguistic knowledge, targeted data, careful evaluation, and human expertise.
The future will not simply be a more powerful version of the present. As answers become easier to generate, good questions, deep understanding, and judgment become more valuable. Our challenge is not merely to build more powerful genies, but to choose our lamps wisely, and become better wish-makers.

Bio: Nizar Habash is a Professor of Computer Science at New York University Abu Dhabi (NYUAD). He is also the director of the Computational Approaches to Modeling Language (CAMeL) Lab. Professor Habash specializes in natural language processing and computational linguistics. He received his PhD in Computer Science from the University of Maryland College Park in 2003. His research includes extensive work on machine translation, morphological analysis, and computational modeling of Arabic and its dialects. Professor Habash has been a principal investigator or co-investigator on over 30 research grants. And he has over 300 publications including a book entitled "Introduction to Arabic Natural Language Processing". Professor Habash is one of the inaugural recipients of the King Salman Academy for Arabic Language Award (2022); he is the recipient of the Antonio Zampolli Prize (2024); and he was selected as a Fellow of the Association for Computational Linguistics (2025). His website is www.nizarhabash.com.

Table of Contents

YUE-PUB: A Cantonese Pragmatic Understanding Benchmark
Chengyan Wu, Ziwei Gong, Jinzhi Chen, Mengmeng Su, Yun Xue and Bolei Ma 1

Beyond Bilingual Transfer: Multilingual Code-Switching in Instruction Tuning
Shunta Asano, Jeonghun Baek and Toshihiko Yamasaki 23

ECQL: An Executable Benchmark for Auditable Corpus Query Semantic Parsing
Yuxi Zheng, Yeqing Liu, Ziming Zhou, Xiaodong Wang, Yongwei Zhang and Jingsong Yu . . . 35

Investigating Learner-Aware Design of LLM-Generated Educational Feedback
Momoka Furuhashi, Kouta Nakayama, Noboru Kawai, Takashi Kodama, Saku Sugawara and Kyosuke Takami 45

Extended-SVD Low-Rank Adaptation: PLORA for Parameter-Efficient Fine-Tuning
Ratan Kumar Gulipilli 74

Improving the Evaluation of Biomedical Question Answering with Few Human Preference Signals
Shakila Mahjabin Tonni, Xiang Dai, Sarvnaz Karimi, Xiaomei Li, Samuel C. Andrew and Stephen Wan 94

Beyond Factual Accuracy: A Benchmark for Evaluating Perspective Alignment in LLMs
Sunghoon Kim and Sangyeop Kim 116

Multi-Signal Interpretability Identifies Task-Critical Circuits for Efficient Model Compression
Srikar Vardhan Mangadoddi 140

Multilinguality after Quantization: Low-Bits, Calibration Languages, Mixed-Precisions, and Cross-Linguality
Muhammad Rafif Priyo Waskito, Alfan Farizki Wicaksono, Alham Fikri Aji and Haryo Akbarianto Wibowo 155

A Systematic Analysis of Linguistic Features in AI-Generated English Text Detection Across Domains and Models
Yassir El Attar, Esra Dönmez, Maximilian Maurer and Agnieszka Falenska 181

Can LLMs Write Reliable Rubrics? A Meta-Evaluation for Experiment Reproduction
Hanhua Hong, Yizhi LI, Jiaoyan Chen, Luu Gia Huy, Sophia Ananiadou, Jung-jae Kim and Chenghua Lin 209

Are Language Models Script-Aware?
David Kletz, Sandra Mitrovic, Ljiljana Dolamic and Fabio Rinaldi 227

PAL-UI: Planning with Active Look-back for Vision-Based GUI Agents
Zikang Liu, Junyi Li, Xin Zhao, Dawei Gao, Yaliang Li and Ji-Rong Wen 244

A Character-Level Neural Approach to Sinhala Sandhi Splitting
Yasas Ekanayaka and Deshan Sumanathilaka 257

WAON: A Large-Scale Japanese Image-Text Dataset for Cultural Adaptation in Contrastive Vision-Language Models
Issa Sugiura, Shuhei Kurita, Yusuke Oda, Daisuke Kawahara, Yasuo Okabe and Naoaki Okazaki 267

Grounded Revision vs. Prior Injection: Probing Retrieval-Augmented Patent Claim Amendment
Josepha Michiko Leo, Hyun-seok Min, Yehoon Jang, Irvan Zidny, Jin-Woo Chung and Sungchul Choi 281

PLAUSIEVE:A Distribution Based Framework for Measuring Model Belief Beyond Logits
Arghodeep Nandi and Tanmoy Chakraborty 300

Thinking Silently for Patent Claim Generation
Lekang Jiang, Wenjun Sun and Stefan Goetz 329

Judging the judge: Can LLMs track attitude-holders in predicates of personal taste?
Yan Cong 342

Useful Diversity or Hallucinated Variety? Persona-Conditioned Synthetic Query Generation for E-Commerce Search
Filippos Ventirozos, Tom Armitage, Mahmoud Oshagh, Peter A. Appleby and Matthew Shardlow 350

Think When Unsure: Leveraging Model Confidence to Decide When to Use Chain-of-Thought
Samuel Lewis-Lim, Xingwei Tan, Zhixue Zhao and Nikolaos Aletras 369

Improving Multimodal Contrastive Learning of Sentence Embeddings with Object-Phrase Alignment
Kaiyan Zhao, Zhongtao Miao and Yoshimasa Tsuruoka 397

CIL: Causal-Intervention and LLM-Synergized Multimodal Fake News Detection
Hongjie Wang, Hongyu Chen and Peilin Jiang 405

Does Continued Pretraining on a Learner Corpus Improve Automated Essay Scoring on English Proficiency Tests? Evidence from EFCAMDAT
Duy Anh Nguyen 415

MuSe-MI: A Framework for Assessing Multi-Session Motivational Interviewing with AI-Generated Patients
Topaz Freizeit, Moreah Zisquit, Kai Golan Hashiloni, Chaitze Tavens, Doron Friedman and Kfir Bar 432

Do LLMs Understand Context? A Knowledge Graph-Based Evaluation Framework
Subavarshana Arumugam, Mamta Nallaretnam, Kithuni Wickramasinghe, Chamath Gunapala, Pragatheeswaran Vipulanandan, Kamal Premaratne and Uthayasanker Thayasivam 479

SelfGrader: LLM Jailbreak Detection via Anchored Token-Level Logits
Zikai Zhang, Rui Hu, Olivera Kotevska and Jiahao Xu 493

Self-Routing RAG: Binding Selective Retrieval with Knowledge Verbalization
Di Wu, Jia-Chen Gu, Kai-Wei Chang and Nanyun Peng 524

Towards Low-Resource Manchu Text Generation: Corpus Construction and Multi-Granularity Augmentation
Meiqi Wang, Zhiyong Xu, Xiaoxin Sun, Ruixin Yu, Zhiyong Guo, Zifan Wang, Shuqi You, Siyu Liu, Xiayu Qin and Suhua Wang 546

PeMA-KGQA: A Multi-Agent LLM Framework for Persian Knowledge Graph Question Answering
Kimia Hosseini, Mohammad Hosseini and Saeedeh Momtazi 560

Code-Prompt: Evaluating Code-LLMs for Natural Language Tasks via Code Aligned Prompts
Yuxuan Chen, Philippe Thomas, Robert Oehring, Felix Krenzien, Sebastian Möller and Roland Roller 580

Evaluating Multi-Dimensional Generalization of Large Language Models in Temporal Extraction Tasks
Fahmid Shahriar Iqbal, Ritam Dutt, Soumitra Das, Arnav Verma and Sagnik Ray Choudhury 605

Mechanistically Understanding and Mitigating the Anchoring Effect in LLMs via Attention Head Intervention
Yoshiki Takenami, Yin Jou Huang, Yugo Murawaki, Koh Takeuchi and Chenhui Chu 630

Beyond Top Words: MonoTM for Topic Modeling with Interpretable Monosemantic Features
Una Joh and Bei Yu 645

Measuring Iterative Temporal Reasoning with Time Puzzles
Zhengxiang Wang and Zeyu Dong 683

Anthropogenic Regional Adaptation in Multimodal Vision-Language Model
Samuel Cahyawijaya, Peerat Limkonchotiwat, Tack Hwa Wong, Carlos Rafael Catalan, Manuel Antonio Rufino, Hitesh Laxmichand Patel, Amit Agarwal, Muhammad Reza Qorib, Vicky Feliren, Holy Lovenia, Aye Hninn Khine, Frederikus Hudi, David Anugraha, Alham Fikri Aji, Romrawin Chumpu, Viet Thanh Pham, Minghan Wang, Mohamed Fazli Mohamed Imam, Ruochen Zhang, Joseph Marvin Imperial, Khumaisa Nur'aini, Do Xuan Long, Musa Izzanardi Wijanarko, Joel Ruben Antony Moniz, Patrick Amadeus Irawan, Hanif Muhammad Zhafran, Isaiah Edri W. Flores, Salsabila Zahirah Pranida, Jun Kevin, Jostin Jerico Rosal, Patricia Nicole Monderin, Kun Kerdthaisong, Ahmad Mustafid, My Chiffon Nguyen, Natchapon Jongwiriyanurak, Siva Worajitwannakul, Haochen Li, Adrian Xuan Wei Lim, Bin Wang, Muhammad Ravi Shulthan Habibi, Lynnette Hui Xian Ng, Mithil Bangera, Yeshil Bangera, Priyaranjan Pattnayak, Dun Li Chan, Sherissa Caren Djuniwar, chochanmyeioo@gmail.com chochanmyeioo@gmail.com and Ming Shan Hee 695

Right for the Wrong Reasons: Investigating the Unrobust Nature of High Accuracy Models in Discourse Relation Prediction
Yi Fan and Wei Liu 713

RabakBench: Evaluating Safety Guardrails in Singapore's Multilingual Vernaculars
Gabriel Chua, Leanne Tan, Ziyu Ge and Roy Ka-Wei Lee 729

Understanding Knowledge Distillation in Post-Training: When It Helps and When It Fails
Xin Liu, Simin Ma, Shujian Liu, Song Wang, Sathish Reddy Indurthi, Haoyun Deng, Lu Wang and Kaiqiang Song 754

Two-Stage LLM-as-a-Judge Using Required and Bonus Criteria
Kohei Tsunefuji, Akihiro Tamura, Daichi Nakatsuji, Hiroaki Oobayashi, Mitsuo Hamada and Tsuneo Kato 768

Stylistic Evolution and LLM Neutrality in Singlish Language
Linus Tze En Foo, Weihan Angela Ng, Wenkai Li and Lynnette Hui Xian Ng 778

Linguistic Cues in Political Debates: Lexical Choices and Evaluative Framing in GPT-Generated Commentary
Yu-Yu Chang, Chung-Chi Chen and Ming-Hung Wang 791

Reasoning under Vision: Understanding Visual-Spatial Cognition in Vision-Language Models for CAPTCHA
Python Song, Luke Tenyi Chang, Yun-Yun Tsai, Penghui Li and Junfeng Yang 803

VMMU: A Vietnamese Multitask Multimodal Understanding and Reasoning Benchmark
Vy Tuong Dang, An Vo, Emilio Villa-Cueva, Quang Tau, Duc Dm, Thamar Solorio and Daeyoung Kim 825

Mirrored Frames: LLM-Based Narrative Framing Analysis of Thai-Cambodian Border Conflict Coverage
Siwakorn Bubphasawan and Attapol Rutherford 871

HiCAP: Hierarchical Cultural Adaptation via Personality for Dialogue Generation
Lingzhi Shen, Xiaohao Cai, Yunfei Long, Imran Razzak and Shoaib Jameel 885

PrivacyEvalClaw: A Tool-Augmented Deliberative Agent for Auditing Privacy Leakage in LLM Jailbreak Responses
Krishna Kanth Nakka, Yujun Wang, Shijie Zhou, Tu Nguyen, Gabriele Gelardi and Xuebing Zhou 900

Hidden Consensus: Preference-Validity Compression in Human Feedback
Dorcas Chia Ern Chua, Karen Myn Hui Lee, Jia Yue Tan, Zhen Xue Gue, Han Ying Lim, Norzalena Abdul Hamid, Azima Binti Azmi, Keat Mei Yeong, Aizat Izyani Binti Mujab, Hafsah Noor Azam, Chee Guo Khoo and Chee Seng Chan 937

EffComp: Efficient Prompt Compression via Hybrid Reinforcement & Supervised Learning
Romtham Tungsununtham, Piyawat Lertvittayakumjorn, Atiwong Suchato and Peerapon Vateekul 969

Replacing the Human in the Loop: Towards Automated Evaluation of LLM Mathematical Proof Generalization
Carolina Dias-Alexiou, Edison Marrese-Taylor, Alfredo Solano and Yutaka Matsuo 996

STEERLSN: Training-Free Representation Alignment through Language-Specific Neuron Subspace Interpolation
Minjun Kim, Dongyeon Kim, Hoyun Song, Jihyo Kim and KyungTae Lim 1005

BreakFun: Jailbreaking LLMs via Object Instantiation under Simulated Code Execution
Amirkia Rafiei Oskooei and Mehmet Aktas 1019

Beyond the Login: Continuous Behavioral Authorship Verification for LLM Agent Sessions
Muskan Sharma and Seba Susan 1039

Synthesizing Dialogues from Social Media Data for Mental Health Support: A Case Study on Gambling Disorders
Zhouqing Zhang, Lis Pereira, Wan Jou She, Benjamin Adams, tham.piumsomboon@canterbury.ac.nz tham.piumsomboon@canterbury.ac.nz and Kongmeng Liew 1055

Beyond SFT: Reinforcement Learning for Safer Large Reasoning Models with Better Reasoning Ability
Jinghan Jia, Nathalie Baracaldo and Sijia Liu 1074

Towards the Roof of the World: NLP Progress in Tibetan Languages
Xi Cao, Yijie Li, Yuan Sun, Shiwei Dong, Zeguang Huang and Nuo Qun 1088

Agent Memory with Episodic Retrieval for Financial Decision-Making
Nuoyue Xu, Jiang Liu, Wenxuan Huang, Xiang Zhang, Juntai Cao and Jiaqi Wei 1120

ProcrastiCoT: A Dual-Encoder Cross-Attention Framework for Multi-Dimensional Procrastination Motives Assessment
Chia-Ni Chiang, Chia-Yun Lee, Hsin-Hsuan Chen and Jheng-Long Wu 1149

Position: Real-World Clinical NLP Robustness Requires Messy, Multimodal, Longitudinal, Privacy-Preserving Corpora
Azmine Toushik Wasi and Shahriyar Zaman Ridoy 1163

When Tool-Backed Skill Retrieval Fails: Source-Style Collapse in Executable Capability Retrieval
Yiqi Liu, Joseph James, Yang Wang, Chenghao Xiao and Chenghua Lin 1181

Feedback as an Interface: Controlling Form and Detail in LLM Translation Refinement
Yiqi Liu, Flavjo Tomori, Nafise Sadat Moosavi, Steffen Eger and Chenghua Lin 1204

Greedy Matching Aggregation for Soft Contrastive Learning in Multi-Label Text Classification
Ryoma Kobayashi and Jiyi Li 1222

IBPS: Indian Bail Prediction System
Uddeshya Raj, Puspesh Kumar Srivastava, Praveen Patel, Shubham Kumar Nigam, Noel Shallum and Arnab Bhattacharya 1236

When Reject Turns into Accept: Quantifying the Vulnerability of LLM-Based Scientific Reviewers to Indirect Prompt Injection
Devanshu Sahoo, Manish Prasad, Vasudev Majhi, Jahnvi Singh, Vinay Chamola, Yash Sinha, Murari Mandal and Dhruv Kumar 1261

Dependency-Aware Trajectory Refinement for Efficient Multi-Turn Agent Fine-Tuning
Zhuo Chen, Zhen Zhang, Xinyu Wang and Kewei Tu 1287

ReThink: Adversarially Robust Arithmetic Reasoning via Neuro-Symbolic Soft Constraints
Soikat Das, Sparsh Mittal and Rekha Singhal 1303

Enhancing Dynamic Text-Attributed Graph Modeling via Instruction Tuning
Weili Shi, Haoming Li and Sheng Li 1314

The Social Cost of Intelligence: Emergence, Propagation, and Amplification of Stereotypical Bias in Multi-Agent Systems
Thi-Nhung Nguyen, Linhao Luo, Amardeep Kaur, Rollin Omari, Tamas Abraham, Junae Kim, Thuy-Trang Vu and Dinh Phung 1330

How Good Are LLM Agents? Benchmarking Opinion Dynamics Against Indonesian Electoral Discourse on Social Media
Musa Izzanardi Wijanarko and Derry Tanti Wijaya 1354

From Early Encoding to Late Suppression: Interpreting LLMs on Character Counting Tasks
Ayan Datta, Mounika Marreddy, Alexander Mehler, Zhixue Zhao and Radhika Mamidi 1370

Measuring the Impact of Big Five Personality Traits on Suicide Severity Level Detection with Large Language Models
Kit May Shum, Yusuke Sakai, Masahiro Kaneko and Taro Watanabe 1392

SignSim: A Landmark Driven Model for Phonological Sign Similarity Prediction
Abhishek Bharadwaj Varanasi, Manjira Sinha, Tirthankar Dasgupta and Dr. Charudatta Jadhav 1412

BioCAT5: Concept Mention-Aware Text-to-Text Transfer Transformer for Biomedical Literature
Mohammad Golam Sohrab 1425

Plan–Write–Refine: An Iterative Framework for Argumentative Essay Generation
Zhijun Zhang, Qianlong Wang, Jianzhu Bao, Keyang Ding, Genan Dai, Bowen Zhang, Ruifeng Xu and Yongsheng Liang 1440

Adversarial Prompt Distillation: Efficient Jailbreak Attacks from Large to Small Language Models
Xiang Li, Chong Zhang, Jia Wang, Fangyu Wu, Yushi Li and Xiaobo Jin 1461

RESCUE-BENCH: Towards Relation-Aware Multi-Party Emotional Support Conversation Systems
Haichuan Hu, Yang Xiao, Mingni Tang, Jiawen Duan, Quanjun Zhang, Congqing He, Hao Zhang, Jiashuo Wang, Johan F. Hoorn and Wenjie Li 1475

GuideWeb: A Benchmark for Automatic In-App Guide Generation on Real-World Web UIs
Chengguang Gan, Yunhao Liang, Yoshihiro Tsujii, Tatsunori Mori, Shiwen Ni and Hiroki Itoh 1519

Audio Editing in the Era of Foundation Models: A Survey
Changhao Pan, YiFei Fan, Fan Zhuo, Yifu Chen, Wenxiang Guo, Yu Zhang, Zhiyuan Zhu, Rui Yang, Shengpeng Ji, Chenyuhao Wen, Jiayang Xu, Ke Lei, Xiaoda Yang, Jingyu Lu and Zhou Zhao 1536

Beyond Accuracy: ARIA-Rubrics for Evaluating Audio Reasoning in Large Audio Language Models
Yupei Li, Qiyang Sun, Mohamed Mady, Chenxi Wang, Zhengwei Gong, Berrak Sisman and Björn Schuller 1560

Large Language and Vision-Language Models as World Models: A Survey
Azmine Toushik Wasi, Sadia Tasnim Meem, Riashat Islam and Md Rizwan Parvez 1582

Can LLMs Threaten Human Survival? Benchmarking Potential Existential Threats from LLMs via Prefix Completion
Yu Cui, Yifei Liu, Hang Fu, Sicheng Pan, Haibin Zhang, Cong Zuo and Licheng Wang 1615

Is There a Dr. House in the Machine? A Survey of Clinical and Diagnostic Reasoning in LLM-based Medical Agents
Azmine Toushik Wasi, Abdur Rahman, Riashat Islam and Md Rizwan Parvez 1632

When Choices Become Risks: Safety Failures of Large Language Models under Multiple-Choice Constraints
Yuheng Chen, Zhiyu Wu, Bowen Cheng, Yihang Wu and Tetsuro Takahashi 1668

RegMix-D: Dynamic Data Mixing via Proxy Training Trajectories
Kaiyan Zhao, Zhongtao Miao, Akiko Aizawa and Yoshimasa Tsuruoka 1698

CCTU: A Benchmark for Tool Use under Complex Constraints
Junjie Ye, Guoqiang Zhang, Wenjie Fu, Zelin Li, Tao Gui, Qi Zhang and Xuanjing Huang . . 1708

LOTUS: Language-Vision Optimal Transport for Unified Report Synthesis
Xuhui Li and Jiachen Liu 1746

From Lexical to Entity-Level: Predicting Individual Knowledge Boundaries via Knowledge Graph Label Propagation
Shuyi Yu, Yusuke Sakai, Hidetaka Kamigaito and Taro Watanabe 1763

ChartHal: A Fine-grained Framework Evaluating Hallucination of Large Vision Language Models in Chart Understanding
Xingqi Wang, Yiming Cui, Xin Yao, Shijin Wang, Guoping Hu and Xiaoyu Qin 1779

Do LLMs Respect Concept Boundaries? A Bilingual Diagnostic Benchmark
Renjun Lai and Yixuan Si 1797

FACTA: Structured Fact Alignment for Multi-Hop Retrieval-Augmented Generation
Jiahao Li, Yongxiu Xu, Xumeng Zhang, Hongbo Xu, Minghao Tang, Yubin Wang and Gaopeng Gou 1815

AGOPNs: Null-Space Activation Steering for Jailbreak Defense via Average Gradient Outer Product Concept Extraction
Minh-Nguyen Bui, Binh-Phuong Phan, Thanh-Khang Doan, Bao Huynh-Gia and Phung-Anh Nguyen 1832

BHAID: A Dataset of LLM-Generated Text in Indian Languages with a Dual Contrastive Learning Framework for Detection
Shifali Agrahari, Sujit Kumar and Ranbir Singh Sanasam 1850

LecTrans: A Multimodal Translation Benchmark for Academic Lectures
Frederikus Hudi, Nina Nanakorn, Justin Vasselli, Yusuke Sakai, Hidetaka Kamigaito, Akihiro Tamura and Taro Watanabe 1872

Predicting Emerging Topics from Outliers: A Prospective Study of Weak Signals in Embedding Space
Evangelia Zve, Gauvain Bourgne and Jean-Gabriel Ganascia 1898

PhaseLoRA: Role-Specialised Low-Rank Adaptation through Prompt-Completion Phase Splitting
Rishi Ravikumar and Riza Batista-Navarro 1914

Can Model Editing Survive Fine-Tuning? Evaluating Robustness and Persistence in Text-to-Image Diffusion Models
Feng He, Marco Valentino and Zhixue Zhao 1926

Neuron-Level Emotion Control in Speech-Generative Large Audio-Language Models
Xiutian Zhao, Ismail Rasim Ulgen, Philipp Koehn, Björn Schuller and Berrak Sisman 1947

Explore-over-Graph: Hybrid Embedding–LLM Reasoning for Knowledge Graph Question Answering under Incompleteness
Ola El Khatib and Djellel Difallah 1962

Iterative Topic Taxonomy Induction with LLMs: A Case Study of Electoral Advertising
Alexander Brady and Tunazzina Islam 1976

TokenMapper: A Step Toward Interoperable Speech Token Translation
Tal Kozakov, Tal Rosenwein and Eliya Nachmani 1991

Modeling Triplet Interactions Beyond Pairwise Attention
Mafizur Rahman and Lijun Qian 2010

A Multimodal Text- and Graph-Based Approach for Open-Domain Event Extraction from Documents
Praval Sharma 2027

Scalable Methods for Long-Range Argumentation in Distributed Discourse
Alina Littek and John Lawrence 2042

STEVE: Stabilizing Textual Gradient-Based Prompt Optimization via Error-Driven Refinement and Regularized Verification
Yifan Xu, Yixuan Li, Xinzhuo Li, Yixin Gu, Yifan Shen, Lijun Yu and Haohan Wang 2057

DunbaaBERT: From Sacrifice to Semantics
Iffat Maab, Waleed Jamil and Raphael Schmitt 2074

CLUE: A Contrastive Learning Framework for Unsupervised Event Type Induction and Annotation
Pratik Maitra, Lauren Kwon and Qi Li 2094

The Validity–Fidelity Gap: A Controlled Study of Representation Drift in Structured Extraction Under Redundant Contexts
Gokul Srinath Seetha Ram, Anirudh Vasudevan and Rashmi Elavazhagan 2112

Retrieval-Induced Mirage: When Similar Cases Become Fabricated Evidence in Medical VLMs
Xiaoxu Ma and Xiangbo Zhang 2127

TileBench: A Controlled Benchmark for Performance Evaluation and Bottleneck Diagnosis of Tile-Based Programming Models
Bowen Cui, Zhongchun Zhou, Hao Wu, Tejas Ramesh, Junyu Yin, Jialiang Gu and Keren Zhou 2145

MF-QAT: Multi-Format Quantization-Aware Training for Elastic Inference
Zifei Xu, Sayeh Sharify and Hesham Mostafa 2168

Bi-Phase Training: Learning Efficiently From Scratch
Abel Gurung and Joseph Campbell 2189

NEXUS: Structured Runtime Safety for Tool-Using LLM Agent
Elias Hossain, Md Mehedi Hasan Nipu, Tasfia Nuzhat Ornee, Rajib Rana and Niloofar Yousefi 2203

SFMP: Fine-Grained, Hardware-Friendly and Solver-Free Mixed-Precision Quantization for Large Language Models
Xin Nie, Haicheng Zhang, Liang Dong, Beining Feng, Jinhong Weng and Guiling Sun 2233

Markhor: A framework for Adapting LLMs to Urdu via Alignment-Free Tokenizer Replacement
Muhammad Zain Ali, Yuxia Wang, Muhammad Arslan Manzoor, Tony C Smith and Bernhard Pfahringer 2260

HakushoBench: A Japanese Chart and Table VQA Benchmark from Governmental White Papers
Issa Sugiura, Shuhei Kurita, Yusuke Oda and Naoaki Okazaki 2273

YFPO: Yoked Feature Preference Optimization with Neuron-Guided Rewards
Yifan Le 2291

Disentangling Perception and Reasoning in Multimodal LLMs via Reward Design
Omar Sharif, Eftekhar Hossain, Nikhil Singh and Patrick Ng 2299

Scale-Aligned Reinforcement Learning for Adaptive Computation Allocation in Reasoning Language Models
Yilin Liu, Xuan Wang, Zheng Yang, Zhuoran Li, Fan Wan, Wu Chen and Xiaojie Wang 2323

UniGuardian: A Unified Defense for Detecting Prompt Injection, Backdoor Attacks and Adversarial Attacks in Large Language Models
Huawei Lin, Yingjie Lao, Tony Geng, Tan YU and Weijie Zhao 2338

ProgMem: Programmatic Belief-State Memory for Long-Horizon Conversational Agents
Pengchao Zheng, Xinan Wang, Ziyi Wei, Haoran Liu, Lei Wei, Jiajun Niu, Bohan Guo and Han Wu 2363

BAEmpEval: A Framework for Probing Large Audio Language Models on Multidimensional Empathetic Response Assessment
Jingxi Liang, Pengze Guo, Zhiwen Xie, Qifeng Wang, Yujia Huo and Derek F. Wong 2377

SQLStructEval: Structural Evaluation of LLM Text-to-SQL Generation
Yixi Zhou, Fan Zhang, Zhiqiao Guo, YU Chen, Haipeng Zhang, Preslav Nakov and Zhuohan Xie 2404

NüshuVoice: Reviving the Voice of Endangered Nüshu with Pitch-Aware Text-to-Speech
Hongkun Yang, YI Xinhui, Zhao Xiyan, Yibo Meng, Lionel Z. Wang, Lixu Wang, Yaqi Zhang, Ruiqi Chen, Xuanyue Zhao, Lanxin Zhang, Yu Zeng, Chu Weijia, Yiming Ma, Chenyu Liu, Jianghao Lin and Xu Xin 2423

GenState-AI: State-Aware Dataset for Text-to-Video Retrieval on AI-Generated Videos
Minghan Li, Tongna Chen, Tianrui Lv, Yishuai Zhang, Suchao An and Guodong Zhou 2435

Quantifying Uncertainty Through Coverage in Latent Semantic Space
Gabriele Carrino, Brown Ebouky, Katja Hose and Mattia Rigotti 2449

Selective Reflection for Context-Sensitive Emotion Recognition in Conversation
Xinyu Wang, Weijie Feng, Hongchuang Wang and Zhiyong Cheng 2466

Style-Aware Translation for Multi-Word Expressions in Low-Resource Languages
Antonios Dimakis and Antonios Anastasopoulos 2482

ReleaseEval: A Benchmark for Evaluating Language Models in Automated Release Note Generation
Qianru Meng, Zhaochun Ren and Joost Visser 2514

EXAM^2: Extending Audio Understanding in Multilingual and Multimodal Analysis
Jiawen Wang, Xiaoxue Gao, Zi Haur Pang and Nancy F. Chen 2527

Student-Centered AI Requires Auditable Evidence of Learning
Xu Aofeng and Sichao Li 2544

The Image Reconstruction Game: Drawing Common Ground Through Iterative Multimodal Dialogue
Sherzod Hakimov, Mattia D'Agostini, Ivan Samodelkin and David Schlangen 2561

Pingala: Prosody-Aware Decoding for Sanskrit Poetry Generation
Manoj Balaji Jagadeeshan, Atul Singh, Nallani Chakravartula Sahith, Amrith Krishna and Pawan Goyal 2584

SMARTS: A Small-to-large Model Coordinated Retrieval and Response Framework for Task-Oriented Dialogue
Jiale Chen, Xuelian Dong, Shengcheng Ye, Wenxiu Xie, Kun Zeng and Tianyong Hao 2603

Activation Residual Hessian Quantization for Low-Bit Large Language Model Inference
YiFeng Wang, Zhun Sun, Hakaze Cho, Linlin Wang and Keisuke Sakaguchi 2613

Beyond Correctness: A Framework for Analyzing Reasoning and Faithfulness of Multi-hop Question Answering
Jiyuan Liu, Yi Xie, Yanghui Rao, Fu Lee Wang, Jianxing Yu and Li Qing 2628

Prompt Rewrites Are Semantic Actions: Testing Whether Language Models Obey Typed Operator Laws
Zida Yang, Weiqing Zeng, Guquan You and Yu Zheng 2649

When Quality Filters Backfire: Agreement-Based Latent Filtering Hurts Bootstrap Pretraining
Suan Lee and JaeSeong Kim 2672

A Survey of Neurosymbolic Approaches to Propaganda and Persuasion Technique Detection
Alysa Ziying Tan and Daniel Kainovan Handoyo 2688

FairJudge: Abstention-Aware Multimodal Judges for Fairness and Alignment Evaluation in Text-to-Image Models
Zahraa Al Sahili, Maimuna Nowaz, Maryam Fetanat, Ioannis Patras and Matthew Purver . . . 2700

WK-Drift: Benchmarking Procedural Memory Maintenance under World Knowledge Drift
Pratyay Banerjee, Masud Moshtaghi and Ankit Chadha 2719

DeScribe: Discourse-Driven Extraction and Structuring of Claim Verification for Better Faithfulness Evaluation
Yang Zhong, Yang Janet Liu and Diane Litman 2737

An Empirical Study of Many-Shot In-Context Learning for Machine Translation of Low-Resource Languages
Yinhan Lu, Gaganpreet Jhajj, Chen Zhang, Anietie Andy and David Ifeoluwa Adelani 2772

Bridging the Long-Tail Gap: Robust Retrieval-Augmented Relation Completion via Multi-Stage Paraphrase Infusion
Fahmida Alam, Mihai Surdeanu and Ellen Riloff 2805

DKCD: Domain Knowledge-Enhanced Causal Discovery from Unstructured Data
Xin Li, Jin Li, Shoujin Wang, Kun Yu and Fang Chen 2824

When more supervision hurts: reward design for structured biomedical reasoning in 4B models
Batuhan Nursal, Christophe Ye, Xinyu Guo, Xin He and Cassie S. Mitchell 2851

Modeling Discourse Profiles in Climate Misinformation Discourse
Ge Gao, Zhengyang Shan, Bryan A. Plummer, Margrit Betke and Derry Tanti Wijaya 2867

CULTURETALK-ID: A Multi-Task Dialogue Benchmark for Cultural Commonsense in Indonesian Local Languages
Muhammad Dehan Al Kautsar, Salsabila Zahirah Pranida, Bilal Elbouardi and Fajri Koto . . 2888

How Is Automated Research Evaluated? A Survey of Benchmarks and Evaluation Practices
Liulei Zhang, Dejing Zhou, Chuyue Huang, Guanhua Chen, Yutong Yao, Lidia S. Chao, Chi Man Vong and Derek F. Wong 2913

Pretraining Builds Cooperation, Not Components: Tracking Inter-Layer Coordination via Lie Brackets
JaeSeong Kim and Suan Lee 2952

Where Does Plan Validity Live in Grammar-Constrained LLM Planning? An Applicability Mask Induced from Environment Rollouts
Nijesh Upreti 2965

Corrective Retrieval Augmented Generation
Shi-Qi Yan, Jia-Chen Gu, Yun Zhu and Zhen-Hua Ling 2991

AI-Sustainability Discourse Across Social Media: Sentiment, Topics, and Engagement on LinkedIn, X, and Bluesky
Kholoud Khalil Aldous, Md. Rafiul Biswas, Mabrouka Bessghaier, Shimaa Ibrahim, Kais Attia and Wajdi Zaghouani 3004

WhisperX-LLM-TermFixer: Enhancing Whisper with LLM-Guided Post-Decoding for Domain-Specific Terminology
Ronghao Pan, José Antonio García-Díaz and Rafael Valencia-García 3019

BlockBench: Contamination-Controlled Evaluation of Large Language Models for Code Analysis
Paul Osemudiame Oamen, Courage Ochuko Mene, Abraham Ugbedeojo Daniel, Ananya Mukherjee, Pius Onobhayedo and Wei Zhao 3043

SAFE-KV: Causal Suffix Repair for Behavior-Preserving Non-Prefix KV Cache Reuse in RAG
Xingjian Zhang 3063

InLegalLaySum: A Novel Benchmark for Lay Summarization of Indian Legal Documents
Trishita Saha, Debtanu Datta, Shubham Tiwari, Mohit Kaushal, Priyank Rathore, Mohammad Khalid Pandit, Pawan Goyal and Saptarshi Ghosh 3099

WISE-Flow: Workflow-Induced Structured Experience for Self-Evolving Conversational Service Agents
Yuqing Zhou, Zhuoer Wang, JK, Hong Wang, Samson Koelle, Ziwei Zhu and Wei Niu 3124

Copy Less, Ground More: Selective and Confidence-Aware Decoding for Grounded Legal Generation
Tan-Minh Nguyen, Truong Dinh Do, Hiep K. V. Nguyen and Le-Minh Nguyen 3144

When Do Supervised UQ Ensembles Improve LLM Hallucination Detection? A Robustness Study
Mohit Singh Chauhan, Vipin Gyanchandani and Dylan Bouchard 3169

FAIR: Balancing Identity-Associated Training Signals to Mitigate Personalization Bias
Anvesh Rao Vijjini and Snigdha Chaturvedi 3191

SAVER: Selective Auditing of Verbal Evidence for Error Recovery in VLM Change Reasoning
Youdi Li 3209

When and How to use In-context Learning for modern LVLMs? An Extensive Investigation
Pengxiang Li, Yu Zhao, Jiaxin Yu, Liang Xie and Erwei Yin 3228

Analyzing Multilingual Parametric Knowledge Retrieval with Sparse Autoencoders
Marc E. Canby, Ikhyun Cho and Julia Hockenmaier 3250

DECK: A Consistency × Confidence Taxonomy of LLM Hallucinations
Mohit Singh Chauhan 3265

DEAF: A Benchmark for Diagnostic Evaluation of Acoustic Faithfulness in Audio Language Models
Jiaqi Xiong, Yunjia Qi, Qi Cao, Yu Zheng, Yutong Zhang, Ziteng Wang, Ruofan Liao, Weisheng Xu and Sichen Liu 3286

Interactive and Scalable Evaluation of LLM Math Tutoring via Simulated Teacher-Student Dialogues
Gihyeon Choi and Harksoo Kim 3300

OMIT the Action: Measuring Framing-Invariant Omission Bias under Philosophical Disagreement
Sihyeon Lee, Jihun Song, Chanwoo Kim, Jiwoo Kum and Chanjun Park 3314

Structured but Silent: Probing Capability Requirements in LLM Hidden States
Kyojun Choo, Minsoo Song, Yunju Kang and Chanjun Park 3335

CASCADE: Cognitive Assessment via Sentence Context and Distortion Extraction
Archit Narumanchi, Azamat Rashidov and Jong C. Park 3376

One Prompt Does Not Fit All: Self-Meta-Evolve for Personalized Information Extraction
Hongliang Li, Lu Wang, Yong Xu, Hanyang Chen, Zhitao Hou, Xiaoting Qin, Song Ge, Qingwei Lin and Dongmei Zhang 3392

Models Know Their Shortcuts: Deployment-Time Shortcut Mitigation
Jiayi Li, Shijie Tang, Gün Kaynar, Shiyi Du and Carl Kingsford 3408

Entailed Opinion Matters: Improving the Fact-Checking Performance of Language Models by Relying on their Entailment Ability
Gaurav Kumar, Ayush Garg, Debajyoti Mazumder, Aditya Kishore, Babu Kumar and Jasabanta Patro 3431

COAT: An Orthogonal Benchmark for Diagnosing Chinese Adversarial Toxicity Detection Failures
Jiaqi Hu, Mingshuai Liu, Haobo Wu and Xin Li 3453

Mass-Tail Attribution Entropy: A Regime-Conditional Decoder for Vision-Language Model Hallucination
Jaegwang Shin and Suan Lee 3466

IndexRAG: Index-Time Reasoning for Multi-Hop Retrieval-Augmented Generation
Zhenghua Bao and Yi Shi 3491

GPC-E4: Generalized Consistency and Entropic Evidence for Backdoor Detection in Vision Language Models
Xuan Wang, Xuehu Yan, Yuliang Lu, Kejiang Ye, Cheng-zhong Xu and Xitong Gao 3507

TripPattern: A Pattern-based Text Watermarking Method for Large Language Models
Sangjun Moon, Dasom Choi, Jingun Kwon, Hidetaka Kamigaito, Taro Watanabe and Manabu Okumura 3523

IMFD: End-to-end Multi-Face Forgery Detection through Instruction-based Large Vision-Language Models
Dasom Choi, Sangjun Moon, Hyeongchan Im, Jaeeon Park, Jingun Kwon, Hidetaka Kamigaito, Taro Watanabe and Manabu Okumura 3539

Efficient Many-Shot In-Context Learning with Task-Level Compression and Instance-Level Retrieval
Masayuki Kawarada, Ukyo Honda and Sho Takase 3547

Beyond EOS: Revealing the Internal Monologue of Language Models
Aishwarya Maheswaran and Maunendra Sankar Desarkar 3556

Polar: A Benchmark for Evaluating Political Bias in LLMs
Sangho Kim, Heejin Kim, Hyunggeun Jeon, Yoonhee Park and Jaejin Lee 3575

Reducing Logprob Without Causal Verification: Diagnosing SAE-Based PII Suppression in LLMs
Chaerin Lee, Wonho Lee, Sohee Park and Daeseon Choi 3636

GRAIN: Gradient-Based Interaction Graph of Neurons for Explainable Misinformation Detection
Rathod Darshan D and Rahul Mishra 3655

SIEVE: Slot-wise Evaluation of Information-Role Retention in Fixed-Length Meeting Summarization
Jumpei Nagasawa and Seiya Shibayama 3671

Encoded but Not Routed: Explaining the Table-Chart Gap in Scientific Claim Verification
Sunisth Kumar, Xanh Ho, Tim Schopf, Andre Greiner-Petter, Florian Boudin and Akiko Aizawa 3700

Chained Cross-Modal Learning for Modality-Agnostic Embeddings
Naoki Tamura and Taiki Sekii 3710

SASU: Structure-Anchored Semantic Unification for Code Search
Qian Zhang, Hao Li and Hao Jiang 3724

EMR: Self-Evolving Medical Multi-Agent System via Experience Mining and Reuse
Dongsheng Shi, Yue Li, Xin Yi and Linlin Wang 3736

Psychometric Benchmark and Bias Assessment in Open-Source Pre-trained NLI Models
Fadi Amun, Ofir Cohen, Aviad Elyashar, Odeya Cohen and Rami Puzis 3753

Not All Graphs Are Maps: A Task-Agnostic Exploration Framework for Mobile GUI Navigation
Adamo Orsini, Seyed Mehdi Ayyoubzadeh, Sadman Rakib Pinon, Hannuo Zhang, Koosha Gholipour Baradari, Yuanhao YU, Xinxin Zuo and Yang Wang 3772

Skin-Deep: A Geometric Diagnostic for Alignment Fragility in Large Language Model Representations
Dongyub Jude Lee, Jungseob Lee, Seungyoon Lee, Seongtae Hong, Suhyune Son, Sugyeong Eo, Jaehyung Seo and Heuiseok Lim 3795

ArabMisinfo: A Decade-Long Arabic Facebook Corpus for Misinformation Discourse and Engagement Analysis
Wajdi Zaghouani, Shimaa Ibrahim, Mabrouka Bessghaier and Md. Rafiul Biswas 3809

DiagBench: Evaluating Diagnostic Revision in Large Language Models with Longitudinal Inpatient Records
Huawei Feng, Dongsheng Shi and Linlin Wang 3821

XChoice: Explainable Evaluation of AI–Human Alignment in LLM-based Constrained Choice Decision Making
Weihong Qi, Fan Huang, Rasika Muralidharan, Jisun An and Haewoon Kwak 3833

Molecule Generation through Reasoning with Large Language Models
Congfeng Cao, Di Wu, Zhi Zhang, Jia-Hong Huang, Jelke Bloem and Khalil Sima'an 3862

Breaking the Speech-Act Contract: Uptake Failures and Void Commitments in Large Language Models
Sahana Varadaraju, Bharathwaj Vijayakumar and Lathish Balaji Baskaran 3880

Towards Robust Numerical Claim Verification
Peter Røysland Aarnes and Vinay Setty 3894

FreakOut-LLM: The Effect of Emotional Stimuli on Safety Alignment
Daniel Kuznetsov, Ofir Cohen, Karin Shistik, Asaf Shabtai and Rami Puzis 3916

Personalization Matters: Long-Horizon Conversation Agent with User-Centric Information in Online Shopping Interactions
Rena Wei Gao, Yue Dai, Hao Guan, Shengxiang Gao, Wangyang Wu, Yixin Shen and Jey Han Lau 3934

From Textual Columns to Query Plans: A Unified Relational-Semantic Execution Framework for Hybrid Query Processing
Nima Shahbazi, Seiji Maekawa, Nikita Bhutani and Estevam Hruschka 3979

Behaviorally Grounded User Profiles from the Wild for Personalized Alignment and Multi-Perspective Reasoning
Yuxuan Li, Victor Zhong and Ehsan Kamalloo 4003

PPL-Factory: Task-Aware and Budget-Aware Data Selection from Language Modeling to Reasoning
Hang Zhang and Warren J. Gross 4025

Training free Cross-Modal Alignment for Multilingual Image Retrieval
Usman Shahid, Manan Jain and Natalie Parde 4039

No Convo! The Illusion of Reliability in Online Toxicity Detection
Rachita Rajesh, Christina A Chance, Ashima Suvarna, Kai-Wei Chang and Saadia Gabriel . 4050

BetXplain: An Explanation-Annotated Dataset for Detecting Manipulative Betting Advertisements on Social Media
Msvpj Sathvik, Parmitha Vangapandu, Nishit Rahul Rane, Sathwik Narkedimilli, Mark Lee and Akrati Saxena 4073

HYDRA: Evaluating Logical Thinking Abilities of LLMs through Formulating Hypotheses and Verification Methods
Shin-nosuke Ishikawa, Hirotsugu Ohba, Masato Todo, Atsushi Yoshino and Masahiro Daimon 4090

GCoTD: A Non-Intrusive Diagnostic Pipeline for Black-Box LLM Backdoor Detection via Graph Chain-of-Thought
Chengyu Jia, AprilPyone MaungMaung, Huy H. Nguyen, Isao Echizen and Jinyin Chen . . . 4106

TokenMix: Efficient Batched Inference for Mixture-of-Experts Models with Joint Expert Selection
Aradhana Mohan Parvathy, Sarada Krithivasan, Swagath Venkataramani, Anand Raghunathan and Viji Srinivasan 4128

From Chains to Graphs: Self-Structured Reasoning for General-Domain LLMs
Yingjian Chen, Haoran Liu, Yinhong Liu, Sherry T. Tong, Aosong Feng, Jinghui Lu, JTzhang, Yusuke Iwasawa, Yutaka Matsuo and Irene Li 4141

CORA: A Protocol for Diagnosing Boundary Robustness in Text-to-Audio Retrieval under Query Reformulations
Jae Min Woo, Kyongmin Kong, Bogyung Jeong, Minjeong Kim, HaeJun Yoo and Du-Seong Chang 4160

It Hears, It Sees too: Multi-Modal LLM for Depression Detection By Integrating Visual Understanding into Audio Language Models
Xiangyu Zhao, Yaling Shen, Yiwen Jiang, Zimu Wang, Jiahe Liu, Maxmartwell H Cheng, Guilherme C Oliveira, Robert Desimone, Dominic Dwyer and Zongyuan Ge 4185

UniToR: Towards Universal Token-level Routing for Efficient Reasoning in Large Language Models
Nguyen-Khang Le, Truong Dinh Do and Le-Minh Nguyen 4197

CuBEs: Culturally-Situated Behavioral Evaluations and the Limitations of Culture-Blind LLM Judges
Hoda Ayad, Tanu Mitra and Abhishek Mukherji 4216

K-FORGE: Fisher-Guided Initialization for Preference-Based LLM Unlearning
Daniil Moskovskiy and Alexander Panchenko 4233

MOOC-MESE: Dataset and Framework for Multimodal Entity Set Expansion in MOOCs
Shichao Yue, Jing Wan and Yubin Chen 4249

Self-Route: Automatic Mode Switching via Capability Estimation for Efficient Reasoning
Yang He, Xiao Ding, Bibo Cai, Yufei Zhang, Kai Xiong, Zhouhao Sun, Bing Qin and Ting Liu 4259

No Masking Needed: Predicting Latent Semantics for Efficient Pre-training of Encoder-based Language Models
Gábor Berend 4271

Long-Term Agent Memory through Complementary Curation and Consolidation
Abhishek Gupta, Dongge Han, Daniel Madrigal, Wei-Ning Chen, Mayukh Das, Victor Rühle and Saravan Rajmohan 4286

RASER: Recoverability-Aware Selective Escalation Router for Multi-Hop Question Answering
Yuyang Li, Zihe Yan and Tobias Käfer 4295

POLAR: Ontology-Guided Risk Prevention for Tool-Calling LLM Agents
Yunju Kang, Seonghyeon Cho, Irene Li, Yeo Chan Yoon and Chanjun Park 4312

Outgrowing the Curse of Multilinguality: Scaling Strategies for Efficient Language Adaptation
Kevin Glocker, Kätriin Kukk, Romina Oji, Marcel Bollmann, Marco Kuhlmann and Jenny Kunz 4329

Vouch: Counterfactual Contracts for Multi-Agent Coordination
Abhay Bhandarkar and Vedanth Nanesha 4345

Towards Smart Exploration: Reinforcement Learning for LLM Agents Guided by Meta-Actions
Hanwen Tong, Zhengqi Zhang, Ruixuan Lu, Jincheng Dai, Bin Yang, Bo Zheng and Feng Wei 4369

Towards Universal Calibration for In-Context Classification
Awik Dhar, Sunita Sarawagi and Arpit Agarwal 4389

Data Augmentation for Improving Contextual Faithfulness Under Knowledge Conflict
Mohith Rajesh, Raj Maheshwari, Vashisth Tiwari, Yonatan Bisk, Emma Strubell, Jeremiah Milbauer and Jared Fernandez 4419

Adaptation, Not Algorithm: LoRA and Full Fine-Tuning Show Comparable Black-Box Jailbreak Degradation in Five Open-Weight LLMs
Taleef Tamsal and Jonathan Rusert 4432

Cultural Awareness is Represented but Not Decoded: Tracing Mythological Knowledge across 18 Open-Source LLMs
Iaroslav Chelombitko and Mika Hämäläinen 4459

Leveraging Articulatory Phonetic Group Recognition for Robust Speech Enhancement for Cochlear Implants in Complex Acoustic Environments
Sohini Saha, Leslie M. Collins and Boyla Mainsah 4503

Superficial Reflection or Genuine Thought? A Fine-Grained Cognitive Analysis of Large Reasoning Models
Yuxiang Chen, Zuohan Wu, Ziwei Wang, Xiangning Yu, Xujia Li, Linyi Yang, Mengyue Yang, Jun Wang and Lei Chen 4530

Beyond Diagnosis: Evaluating Recommendation Stability under Counterfactual Skin-Tone Perturbations
Hyunwoo Yoo, Kyusik Kim and Gail Rosen 4549

Revisiting Co-Evolution for Multi-Turn Safety Alignment: Limitations and Alignment Tax Mitigation
Yoshiaki Sata and Kouta Nakayama 4565

The Cost of Compression: A Rate-Distortion Limit on Factual Hallucination
Wang Xi, Shijia Xu and Rongfeng Guo 4579

Optimizing Reasoning LLMs for Automated Feature Engineering via Multi-Stage Direct Preference Optimization
Mingzhi Hu, Menghai Pan, Yuzhong Chen, Junpeng Wang, Yingtong Dou, Vineeth Rakesh, Hoang Anh Duy Le, Uday Singh Saini, Xin Dai, Xiran Fan, Yuqing Liu and Yiwei Cai 4599

Synthesizing Code Completion Trajectories for Mid-training in Software Engineering
Siwei Xu, Qiyuan Liu and Ruifan Li 4619

Parameter-Efficient Domain Adversarial Training for Lightweight Framewise Phoneme Classification
Sohini Saha, Leslie M. Collins and Boyla Mainsah 4631

$M^4olGen : Multi-Agent, Multi-Stage Molecular Generation under Precise Multi-Property Constraints$
Yizhan Li, Florence Cloutier, Sifan Wu, Ali Parviz, Boris Knyazev, Yan Zhang, Glen Berseth and Bang Liu 4649

When Should Language Models Search? A Topology-Aware Analysis of Reasoning Strategies
Dibyanayan Bandyopadhyay, MD Tanzeel Adam Khan and Asif Ekbal 4663

Evaluating LLM-as-a-Judge Beyond Score Alignment: A Psychometric Analysis of Residual Judging Difficulty
Longwei Cong, Sonja Hahn, Sebastian Gombert, Leon Camus, Fabian Zehner, Hendrik Drachsler and Ulf Kroehne 4682

Efficient Multimodal Planning Agent for Visual Question-Answering
Zhuo Chen, Xinyu Geng, Xinyu Wang, Yong Jiang, Zhen Zhang, Pengjun Xie and Kewei Tu4698

YaPO: Learnable Sparse Activation Steering Vectors for Domain Adaptation
Abdelaziz Bounhar, Rania Elbadry, Hadi Abdine, Preslav Nakov, Michalis Vazirgiannis and Guokan Shang 4719

Evaluation of Multilingual LLMs Personalized Text Generation Capabilities Targeting Groups and Social-Media Platforms
Dominik Macko 4741

Reading Between the Layers: Cross-Modal Alignment and Representational Hierarchy in Vision Transformers
Sugam Karki, Raúl Vázquez and Yova Kementchedjhieva 4761

Large Language Models for Solid-Earth Geoscience: Applications, Modalities, and Memory-Reason-Act Techniques
Yisu Chen, Yiran Zhang, Yifeng Dai, Puzhen Wu, Sirui Li, Wei Liu and Yihao Ding 4776

Probing Reasoning Length Control in Language Models via Neurons
Dongwook Min, Jeonhwi Kang, Cheoneum Park and Seung-Hoon Na 4791

SeEM: Selective Dictionary-based Generation
Hongyuan Lu, Zixuan Li and Wai Lam 4809

Beyond Consensus: Diagnosing Coverage, Majority Bias, and Grounding Equity in High-Stakes Multi-Document Diversity Summarization
Jane Arleth dela Cruz, Iris Hendrickx and Martha Larson 4831

DPDisc: From Factoid Questions to Data Product Requests for Open-World Data Product Discovery over Tables and Text
Liangliang Zhang, Nandana Mihindukulasooriya, Niharika S. D'Souza, Sola Shirai, Sarthak Dash, Yao Ma and Horst Samulowitz 4846

Cleansing Fine-Tuning Data to Mitigate Backdoors in LLMs via the Base Model's Inherent Generative Abilities
Qi Zhao, Yan-Keon Elster and Christian Wressnegger 4866