

Conference on Empirical Methods in Natural Language Processing (EMNLP 2025)

Industry Track

Suzhou, China
4-9 November 2025

Volume 1 of 4

ISBN: 979-8-3313-2867-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>RAVEN++: Pinpointing Fine-Grained Violations in Advertisement Videos with Active Reinforcement Reasoning</i>	
Deyi Ji, Yuekui Yang, Liqun Liu, Peng Shu, Haiyang Wu, Shaogang Tang, Xudong Chen, Shaoping Ma, Tianrun Chen and Lanyun Zhu	1
<i>SAGE: A Generic Framework for LLM Safety Evaluation</i>	
Madhur Jindal, Hari Shrawgi, Parag Agrawal and Sandipan Dandapat	11
<i>CRAB: A Benchmark for Evaluating Curation of Retrieval-Augmented LLMs in Biomedicine</i>	
Hanmeng Zhong, Linqing Chen, Wentao Wu and Weilei Wang	34
<i>VENUS: A VLLM-driven Video Content Discovery System for Real Application Scenarios</i>	
Minyi Zhao, YI Liu, Jianfeng Wen, Boshen Zhang, Hailang Chang, Zhiheng Ouyang, Jie Wang, Wensong he and Shuigeng Zhou	50
<i>FT-MDT: Extracting Decision Trees from Medical Texts via a Novel Low-rank Adaptation Method</i>	
Yuheng Li, Jiechao Gao, Wei Han, Wenwen Ouyang, Wei Zhu and Hui Yi Leong	65
<i>PolyNorm: Few-Shot LLM-Based Text Normalization for Text-to-Speech</i>	
Michel Wong, Ali Alshehri, Sophia Kao and Haotian He	77
<i>Audio Query Handling System with Integrated Expert Models and Contextual Understanding</i>	
Naveen Vakada, Arvind Krishna Sridhar, Yinyi Guo and Erik Visser	86
<i>Generative Reviewer Agents: Scalable Simulacra of Peer Review</i>	
Nicolas Bougie and Narimawa Watanabe	98
<i>Aligning LLMs for Multilingual Consistency in Enterprise Applications</i>	
Amit Agarwal, Hansa Meghwani, Hitesh Laxmichand Patel, Tao Sheng, Sujith Ravi and Dan Roth	117
<i>RCI: A Score for Evaluating Global and Local Reasoning in Multimodal Benchmarks</i>	
Amit Agarwal, Hitesh Laxmichand Patel, Srikant Panda, Hansa Meghwani, Jyotika Singh, Karan Dua, Paul Li, Tao Sheng, Sujith Ravi and Dan Roth	138
<i>LP Data Pipeline: Lightweight, Purpose-driven Data Pipeline for Large Language Models</i>	
Yungi Kim, Hyunsoo Ha, Seonghoon Yang, Sukyung Lee, Jihoo Kim and Chanjun Park	158
<i>Toward Reliable Clinical Coding with Language Models: Verification and Lightweight Adaptation</i>	
Moy Yuan, Han-Chin Shing, Mitch Strong and Chaitanya Shivade	173
<i>Enhancing Talent Search Ranking with Role-Aware Expert Mixtures and LLM-based Fine-Grained Job Descriptions</i>	
Jihang Li, Bing Xu, Zulong Chen, Chuanfei Xu, Minping Chen, Suyu Liu, Ying Zhou and Zeyi Wen	185
<i>PCRI: Measuring Context Robustness in Multimodal Models for Enterprise Applications</i>	
Hitesh Laxmichand Patel, Amit Agarwal, Srikant Panda, Hansa Meghwani, Karan Dua, Paul Li, Tao Sheng, Sujith Ravi and Dan Roth	195
<i>CitySim: Modeling Urban Behaviors and City Dynamics with Large-Scale LLM-Driven Agent Simulation</i>	
Nicolas Bougie and Narimawa Watanabe	215

<i>Evaluating Conversational Agents with Persona-driven User Simulations based on Large Language Models: A Sales Bot Case Study</i>	
Justyna Gromada, Alicja Kasicka, Ewa Komkowska, Lukasz Krajewski, Natalia Krawczyk, Morgan Veyret, Bartosz Przybył, Lina M. Rojas-Barahona and Michał K. Szczerbak	230
<i>Mirror in the Model: Ad Banner Image Generation via Reflective Multi-LLM and Multi-modal Agents</i>	
Zhao Wang, Bowen Chen, Yotaro Shimose, Sota Moriyama, Heng Wang and Shingo Takamatsu	246
<i>Leveraging Product Catalog Patterns for Multilingual E-commerce Product Attribute Prediction</i>	
Bryan Zhang, Suleiman A. Khan and SteCphan Walter	267
<i>ECom-Bench: Can LLM Agent Resolve Real-World E-commerce Customer Support Issues?</i>	
Haoxin Wang, Xianhan Peng, Huang Cheng, Yizhe Huang, Ming Gong, Chenghan Yang, Yang Liu and Jiang Lin	276
<i>ProCut: LLM Prompt Compression via Attribution Estimation</i>	
Zhentao Xu, Fengyi Li, Albert C. Chen and Xiaofeng Wang	285
<i>A Reasoner for Real-World Event Detection: Scaling Reinforcement Learning via Adaptive Perplexity-Aware Sampling Strategy</i>	
Xiaoyun Zhang, Jingqing Ruan, Xing Ma, Yawen Zhu, Jiansong Chen, Ke Zeng and Xunliang Cai	310
<i>Detecting Omissions in LLM-Generated Medical Summaries</i>	
Achir Oukelmoun, Nasredine Semmar, Gaël de Chalendar, Clement cormi Cormi, Mariame Oukelmoun, Eric vibert Vibert and Marc-Antoine Allard	325
<i>LEAF: Learning and Evaluation Augmented by Fact-Checking to Improve Factualness in Large Language Models</i>	
Hieu Tran, Junda Wang, Yujan Ting, Hong yu, Weijing Huang and Terrence Chen	338
<i>ReAct Meets Industrial IoT: Language Agents for Data Access</i>	
James T Rayfield, Shuxin Lin, Nianjun Zhou and Dhaval C Patel	364
<i>ProductAgent: Benchmarking Conversational Product Search Agent with Asking Clarification Questions</i>	
Jingheng Ye, Yong Jiang, Xiaobin Wang, Yinghui Li, Yangning Li, Pengjun Xie and Fei Huang	383
<i>MADS: Multi-Agent Dialogue Simulation for Diverse Persuasion Data Generation</i>	
Mingjin Li, Yu Liu, Huayi Liu, Xiang Ye, Chao Jiang, Hongguang Zhang and Yu Ruan	399
<i>On-device System of Compositional Multi-tasking in Large Language Models</i>	
Ondrej Bohdal, Konstantinos Theodosiadis, Asterios Mpatziakas, Dimitrios Filippidis, Iro Spyrou, Christos Zonios, Anastasios Drosou, Dimosthenis Ioannidis, Kyenghun Lee, Jijoong Moon, Hyeonmok Ko, Mete Ozay and Umberto Michieli	416
<i>Select-then-Route : Taxonomy guided Routing for LLMs</i>	
Soham Shah and Kumar Shridhar	425
<i>FABRIC: Fully-Automated Broad Intent Categorization in E-commerce</i>	
Anna Tigunova, Philipp Schmidt and Damla Ezgi Akcora	442
<i>MKT: A Multi-Stage Knowledge Transfer Framework to Mitigate Catastrophic Forgetting in Multi-Domain Chinese Spelling Correction</i>	
Peng Xing, Yinghui Li, Shirong Ma, Xinnian Liang, Haojing Huang, Yangning Li, Shu-Yu Guo, Hai-Tao Zheng, Wenhao Jiang and Ying Shen	451

<i>End-to-End Aspect-Guided Review Summarization at Scale</i>	
Ilya Boytsov, Vinny DeGenova, Mikhail Balyasin, Joseph Walt, Caitlin Eusden, Marie-Claire Rochat and Margaret Pierson	463
<i>SLOT: Structuring the Output of Large Language Models</i>	
Zhengyuan Shen, Darren Yow-Bang Wang, Soumya Smruti Mishra, Zhichao Xu, Yifei Teng and Haibo Ding	472
<i>QuackIR: Retrieval in DuckDB and Other Relational Database Management Systems</i>	
Yijun Ge, Zijian Chen and Jimmy Lin	492
<i>Benchmarking Deep Search over Heterogeneous Enterprise Data</i>	
Prafulla Kumar Choubey, Xiangyu Peng, Shilpa Bhagavath, Kung-Hsiang Huang, Caiming Xiong and Chien-Sheng Wu	501
<i>RLHF Algorithms Ranked: An Extensive Evaluation Across Diverse Tasks, Rewards, and Hyperparameters</i>	
Lucas Spangher, Rama Kumar Pasumarthi, Nick Masiewicki, William F. Arnold, Aditi Kaushal, Dale Johnson, Peter Grabowski and Eugene Ie	518
<i>Predicting Cross-lingual Trends in Microblogs</i>	
Satoshi Akasaki	530
<i>Generating Fine Details of Entity Interactions</i>	
Xinyi Gu and Jiayuan Mao	540
<i>AutoCVSS: Assessing the Performance of LLMs for Automated Software Vulnerability Scoring</i>	
Davide Sanvito, Giovanni Arriciati, Giuseppe Siracusano, Roberto Bifulco and Michele Carminati	564
<i>SFAL: Semantic-Functional Alignment Scores for Distributional Evaluation of Auto-Interpretability in Sparse Autoencoders</i>	
Fabio Mercorio, Filippo Pallucchini, Daniele Poterti, Antonio Serino and Andrea Seveso	576
<i>Just One is Enough: An Existence-based Alignment Check for Robust Japanese Pronunciation Estimation</i>	
Hayate Nakano and Nobuhiro Kaji	584
<i>Towards Enforcing Company Policy Adherence in Agentic Workflows</i>	
Naama Zwerdling, David Boaz, Ella Rabinovich, Guy Uziel, David Amid and Ateret Anaby Tavor	595
<i>Learning to Translate Ambiguous Terminology by Preference Optimization on Post-Edits</i>	
Nathaniel Berger, Johannes Eschbach-Dymanus, Miriam Exel, Matthias Huck and Stefan Riezler	607
<i>More Data or Better Data? A Critical Analysis of Data Selection and Synthesis for Mathematical Reasoning</i>	
Yike Zhao, Simin Guo, Ziqing Yang, Shifan Han, Dahua Lin and Fei Tan	618
<i>SRS-Stories: Vocabulary-constrained multilingual story generation for language learning</i>	
Wiktor Kamzela, Mateusz Lango and Ondrej Dusek	630
<i>Banking Done Right: Redefining Retail Banking with Language-Centric AI</i>	
Xin Jie Chua, Jeraelyn Ming Li Tan, Jia Xuan Tan, Soon Chang Poh, Yi Xian Goh, Debbie Hui Tian Choong, Foong Chee Mun, Sze Jue Yang and Chee Seng Chan	646

<i>Graph of Attacks with Pruning: Optimizing Stealthy Jailbreak Prompt Generation for Enhanced LLM Content Moderation</i>	
Daniel Schwartz, Dmitriy Beshpalov, Zhe Wang, Ninad Kulkarni and Yanjun Qi	659
<i>Structural Reward Model: Enhancing Interpretability, Efficiency, and Scalability in Reward Modeling</i>	
Xiaoyu Liu, Di Liang, Hongyu Shan, Peiyang Liu, Yonghao Liu, Muling Wu, Yuntao Li, Xianjie Wu, LI Miao, Jiangrong Shen and Minlong Peng	672
<i>Controllable Clustering with LLM-driven Embeddings</i>	
Kerria Pang-Naylor, Shivani Manivasagan, Aitong Zhong, Mehak Garg, Nicholas Mondello, Blake Buckner, Jonathan P. Chang, Khyati Mahajan, Masoud Hashemi and Fabio Casati	686
<i>SpeechLLMs for Large-scale Contextualized Zero-shot Slot Filling</i>	
Kadri Hacioglu, Manjunath K E and Andreas Stolcke	703
<i>NurseLLM: The First Specialized Language Model for Nursing</i>	
Md Tawkat Islam Khondaker, Julia Harrington and Shady Shehata	716
<i>Augmenting Compliance-Guaranteed Customer Service Chatbots: Context-Aware Knowledge Expansion with Large Language Models</i>	
Mengze Hong, Chen Jason Zhang, Di Jiang and Yuanqin He	753
<i>Memory-Efficient Backpropagation for Fine-Tuning LLMs on Resource-Constrained Mobile Devices</i>	
Congzheng Song and Xinyu Tang	766
<i>PrismRAG: Boosting RAG Factuality with Distractor Resilience and Strategized Reasoning</i>	
Mohammad Kachuee, Teja Gollapudi, Minseok Kim, Yin Huang, Kai Sun, Xiao Yang, Jiaqi Wang, Nirav Shah, Yue Liu, Aaron Colak, Anuj Kumar, Wen-tau Yih and Xin Luna Dong	775
<i>Benchmarking LLM Faithfulness in RAG with Evolving Leaderboards</i>	
Manveer Singh Tamber, Forrest Sheng Bao, Chenyu Xu, Ge Luo, Suleman Kazi, minseok Bae, Miaoran Li, Ofer Mendelevitch, Renyi Qu and Jimmy Lin	799
<i>A Multi-Agent Framework for Quantitative Finance : An Application to Portfolio Management Analytics</i>	
Sayani Kundu, Dushyant Sahoo, Victor Li, Jennifer Rabowsky and Amit Varshney	812
<i>Group Preference Alignment: Customizing LLM Responses from In-Situ Conversations Only When Needed</i>	
Ishani Mondal, Jack W. Stokes, Sujay Kumar Jauhar, Longqi Yang, Mengting Wan, Xiaofeng Xu, Xia Song, Jordan Lee Boyd-Graber and Jennifer Neville	825
<i>DASR: Distributed Adaptive Scene Recognition - A Multi-Agent Cloud-Edge Framework for Language-Guided Scene Detection</i>	
Can Cui, Yongkang Liu, Seyhan Ucar, Juntong Peng, Ahmadreza Moradipari, Maryam Khabazi and Ziran Wang	850
<i>Empowering Healthcare Practitioners with Language Models: Structuring Speech Transcripts in Two Real-World Clinical Applications</i>	
Jean-Philippe Corbeil, Asma Ben Abacha, George Michalopoulos, Phillip Swazinna, Miguel Del-Agua, Jerome Tremblay, Akila Jeason Daniel, Cari Bader, Kevin Cho, Pooja Krishnan, Nathan Bodensstab, Thomas Lin, Wenxuan Teng, Francois Beaulieu and Paul Vozila	859
<i>Leveraging the Power of Large Language Models in Entity Linking via Adaptive Routing and Targeted Reasoning</i>	
Yajie Li, Albert Galimov, Mitra Datta Ganapaneni, Pujitha Thejaswi, De Meng, Priyanshu Kumar and Saloni Potdar	871

<i>Can LLMs Narrate Tabular Data? An Evaluation Framework for Natural Language Representations of Text-to-SQL System Outputs</i>	
Jyotika Singh, Weiyi Sun, Amit Agarwal, Viji Krishnamurthy, Yassine Benajiba, Sujith Ravi and Dan Roth	883
<i>Enhancing Foundation Models in Transaction Understanding with LLM-based Sentence Embeddings</i>	
Xiran Fan, Zhimeng Jiang, Chin-Chia Michael Yeh, Yuzhong Chen, Yingtong Dou, Menghai Pan and Yan Zheng	903
<i>Agent vs. Agent: Automated Data Generation and Red-Teaming for Custom Agentic Workflows</i>	
Ninad Kulkarni, Xian Wu, Siddharth Varia and Dmitriy Beshpalov	912
<i>Auto prompting without training labels: An LLM cascade for product quality assessment in e-commerce catalogs</i>	
Soham Satyadharma, Fatemeh Sheikholeslami, Swati Kaul, Aziz Umit Batur and Suleiman A. Khan	937
<i>Harmonizing Diverse Models: A Layer-wise Merging Strategy for Consistent Generation</i>	
Xujun Peng, Anoop Kumar, Jingyu Wu, Parker Glenn and Daben Liu	954
<i>Transparent Reference-free Automated Evaluation of Open-Ended User Survey Responses</i>	
Subin An, Yugyeong Ji, Junyoung Kim, Heejin Kook, Yang Lu and Josh Seltzer	963
<i>Datasets and Recipes for Video Temporal Grounding via Reinforcement Learning</i>	
Ruizhe Chen, Tianze Luo, Zhiting Fan, Heqing Zou, Zhaopeng Feng, Guiyang Xie, Hansheng Zhang, Zhuochen Wang, Zuozhu Liu and Zhang Huaijian	983
<i>SEARA: An Automated Approach for Obtaining Optimal Retrievers</i>	
Zou Yuheng, Wang Yiran Yiran, Tian Yuzhu, Zhu Min and Yanhua Huang	993
<i>UniEDU: Toward Unified and Efficient Large Multimodal Models for Educational Tasks</i>	
Zhendong Chu, Jian Xie, Shen Wang, Zichao Wang and Qingsong Wen	1007
<i>Truth, Trust, and Trouble: Medical AI on the Edge</i>	
Mohammad Anas Azeez, Rafiq Ali, Ebad Shabbir, Zohaib Hasan Siddiqui, Gautam Siddharth Kashyap, Jiechao Gao and Usman Naseem	1017
<i>An Address Intelligence Framework for E-commerce Deliveries</i>	
Gokul Swamy, Aman Gulati, Srinivas Virinchi and Anoop Saladi	1026
<i>LLMs on a Budget? Say HOLA</i>	
Zohaib Hasan Siddiqui, Jiechao Gao, Ebad Shabbir, Mohammad Anas Azeez, Rafiq Ali, Gautam Siddharth Kashyap and Usman Naseem	1035
<i>LLM-Based Dialogue Labeling for Multiturn Adaptive RAG</i>	
Zhiyu Chen, Biancen Xie, Sidarth Srinivasan, Manikandarajan Ramanathan, Rajashekar Maragoud and Qun Liu	1044
<i>RAGulator: Lightweight Out-of-Context Detectors for Grounded Text Generation</i>	
Ian Poey, Jiajun Li1 and Qishuai Zhong	1057
<i>REIC: RAG-Enhanced Intent Classification at Scale</i>	
Ziji Zhang, Michael Yang, Zhiyu Chen, Yingying Zhuang, Shu-Ting Pi, Qun Liu, Rajashekar Maragoud, Vy Nguyen and Anurag Beniwal	1072
<i>Mapping Smarter, Not Harder: A Test-Time Reinforcement Learning Agent That Improve Without Labels or Model Updates</i>	
Wen-Kwang Tsao, Yao-Ching Yu and Chien-Ming Huang	1081

<i>On Assigning Product and Software Codes to Customer Service Requests with Large Language Models</i> Sujatha Das Gollapalli, Mouad Hakam, Mingzhe Du, See-Kiong Ng and Mohammed Hamzeh	1092
<i>Reasoning-Enhanced Domain-Adaptive Pretraining of Multimodal Large Language Models for Short Video Content Governance</i> Zixuan Wang, Yu Sun, Hongwei Wang, Baoyu Jing, Xiang Shen, Xin Dong, Zhuolin Hao, Hongyu Xiong and Yang Song	1104
<i>GSID: Generative Semantic Indexing for E-Commerce Product Understanding</i> Haiyang Yang, Qinye Xie, Qingheng Zhang, Chen Li Yu, Huike Zou, Chengbao Lian, Shuguang Han, Fei Huang, Jufeng Chen and Bo Zheng	1113
<i>Learning from LLM Agents: In-Context Generative Models for Text Casing in E-Commerce Ads</i> Yingxue Zhou, Tan Zhu, Tao Zeng, Zigeng Wang and Wei Shen	1122
<i>Auto-Weighted Group Relative Preference Optimization for Multi-Objective Text Generation Tasks</i> Yuki Ichihara and Yuu Jinnai	1134
<i>Cost-Effective E-Commerce Catalog Translation at Scale Ensuring Named Entity Protection</i> Asier Gutiérrez-Fandiño, Jorge Yero Salazar, Clement Ruin, Alejandro Quintero-Roba, Shangeetha Ravichandran, Jesus Perez-Martin, Pankaj Adsul, Suruchi Garg and Leonardo Lezcano	1148
<i>InstaJudge: Aligning Judgment Bias of LLM-as-Judge with Humans in Industry Applications</i> Myeongjun Erik Jang and Fran Silavong	1158
<i>TelAgentBench: A Multi-faceted Benchmark for Evaluating LLM-based Agents in Telecommunications</i> Sunwoo Lee, Daseong Jang, Dhammiko Arya, Gyoung-eun Han, Injee Song, SaeRom Kim, Sangjin Kim, Seojin Lee, Seokyoung Hong, Sereimony Sek, Seung-Mo Cho, Sohee Park, Sungbin Yoon, Wonbeom Jang and Eric Davis	1173
<i>Taming the Real-world Complexities in CPT E/M Coding with Large Language Models</i> Islam Nassar, Yang Lin, Yuan Jin, Rongxin Zhu, Chang Wei Tan, Zenan Zhai, Nitika Mathur, Thanh Tien Vu, Xu Zhong, Long Duong and Yuan-Fang Li	1212
<i>Classifier-Augmented Generation for Structured Workflow Prediction</i> Thomas Gschwind, Shramona Chakraborty, Nitin Gupta and Sameep Mehta	1227
<i>Efficient and Versatile Model for Multilingual Information Retrieval of Islamic Text: Development and Deployment in Real-World Scenarios</i> Vera Pavlova and Mohammed Makhlof	1239
<i>AutoQual: An LLM Agent for Automated Discovery of Interpretable Features for Review Quality Assessment</i> Xiaochong Lan, Jie Feng, Yinxing Liu, Xinleishi and Yong Li	1250
<i>JSON Whisperer: Efficient JSON Editing with LLMs</i> Sarel Duanis, Asnat Greenstein-Messica and Eliya Habba	1265
<i>L4: Mutual Learning Helps Lifelong Language Learning</i> Jiyong Li, Dilshod Azizov and Shangsong Liang	1275
<i>TTD-SQL: Tree-Guided Token Decoding for Efficient and Schema-Aware SQL Generation</i> Chetan Sharma, Ramasuri Narayanam, Soumyabrata Pal, Kalidas Yeturu, Shiv Kumar Saini and Koyel Mukherjee	1287

<i>Spot the BlindSpots: Systematic Identification and Quantification of Fine-Grained LLM Biases in Contact Center Call Summarization</i>	
Kawin Mayilvaghanan, Siddhant Gupta and Ayush Kumar	1299
<i>HierDiffuse: Progressive Diffusion for Robust Interest Fusion in CTR Prediction</i>	
Ziheng Ni, Congcong Liu, Yuying Chen, Zhiwei Fang, Changping Peng, Zhangang Lin, Ching Law and Jingping Shao	1341
<i>TOBUGraph: Knowledge Graph-Based Retrieval for Enhanced LLM Performance Beyond RAG</i>	
Savini Kashmira, Jayanaka L. Dantanarayana, Joshua Brodsky, Ashish Mahendra, Yiping Kang, Krisztian Flautner, Lingjia Tang and Jason Mars	1349
<i>Thinking with DistilQwen: A Tale of Four Distilled Reasoning and Reward Model Series</i>	
Wenrui Cai, Chengyu Wang, Junbing Yan, Jun Huang and Xiangzhong Fang	1357
<i>Crossing Domains without Labels: Distant Supervision for Term Extraction</i>	
Elena Senger, Yuri Campbell, Rob Van Der Goot and Barbara Plank	1366
<i>I-SEE: An Instruction-tuned, SOP-Enhanced Quality Evaluator for Product Content</i>	
Aniket Joshi, Cyrus Andre DSouza, Sejal Jain, Jitenkumar Babubhai Rana and Promod Yenigalla	1379
<i>Computational Blueprints: Generating Isomorphic Math Problems with Large Language Models</i>	
Jeong-hoon Kim, Jinwoo Nam and Geunsik Jo	1389
<i>Fin-ExBERT: User Intent based Text Extraction in Financial Context using Graph-Augmented BERT and trainable Plugin</i>	
Soumick Sarker and Abhijit Kumar Rai	1402
<i>DecEx-RAG: Boosting Agentic Retrieval-Augmented Generation with Decision and Execution Optimization via Process Supervision</i>	
Yongqi Leng, Yikun Lei, Xikai Liu, Meizhi Zhong, Bojian Xiong, Yurong Zhang, Yan Gao, Yiwu, Yao Hu and Deyi Xiong	1412
<i>FLOW-BENCH: Towards Conversational Generation of Enterprise Workflows</i>	
Evelyn Duesterwald, Siyu Huo, Vatche Isahagian, K. R. Jayaram, Ritesh Kumar, Vinod Muthusamy, Punleuk Oum, Debashish Saha, Gegi Thomas and Praveen Venkateswaran	1426
<i>Format Inertia: A Failure Mechanism of LLMs in Medical Pre-Consultation</i>	
Seungseop Lim, Gibaeg Kim, Wooseok Han, Jean Seo, Hyunkyung Lee, Jaehyo Yoo and Eunho Yang	1437
<i>Extraction of Information Provision Activity Requirements from EU Acquis</i>	
Jakub Piskorski and Dominik Skotarczak	1451
<i>Contrastive Learning Using Graph Embeddings for Domain Adaptation of Language Models in the Process Industry</i>	
Anastasia Zhukova, Jonas Luehrs, Christian Matt and Bela Gipp	1472
<i>From Feedback to Checklists: Grounded Evaluation of AI-Generated Clinical Notes</i>	
Karen Zhou, John Michael Giorgi, Pranav Mani, Peng Xu, Davis Liang and Chenhao Tan ..	1485
<i>FlexDoc: Parameterized Sampling for Diverse Multilingual Synthetic Documents for Training Document Understanding Models</i>	
Karan Dua, Hitesh Laxmichand Patel, Puneet Mittal, Ranjeet Gupta, Amit Agarwal, Praneet Pabolu, Srikant Panda, Hansa Meghwani, Graham Horwood and Fahad Shah	1500

<i>GEMMAS: Graph-based Evaluation Metrics for Multi Agent Systems</i>	
Jisoo Lee, Raeyoung Chang, Dongwook Kwon, Harmanpreet Singh and Nikhil Verma	1522
<i>Knowledge-Augmented Question Error Correction for Chinese Question Answer System with Question-RAG</i>	
Longpeng Qiu, Ting Li, Shuai Mao, Nan Yang and Xiaohui Yan	1533
<i>SMART: Scalable Multilingual Approach for a Robust TOD System</i>	
Karan Malhotra, Arihant Jain, Purav Aggarwal and Anoop Saladi	1543
<i>Think-Search-Patch: A Retrieval-Augmented Reasoning Framework for Repository-Level Code Repair</i>	
Bojian Xiong, Yikun Lei, Xikai Liu, Shaowei Zhang, Pengyun Zhu, Yan Liu, Yongqi Leng, Ling Shi, Meizhi Zhong, Yurong Zhang, Yan Gao, Yiwu, Yao Hu and Deyi Xiong	1555
<i>ASR-EC Benchmark: Evaluating Large Language Models on Chinese ASR Error Correction</i>	
Victor Junqiu Wei, Weicheng Wang, Di Jiang, Yuanfeng Song and Lu Wang	1567
<i>Bidirectional Reasoning Supervision for Multilingual Financial Decision Making</i>	
Muhammad Rafsan Kabir, Jawad Ibn Ahad, Robin Krambroeckers, Silvia Ahmed, M M Lutfelahi, Nabeel Mohammed and Shafin Rahman	1576
<i>Automotive Document Labeling Using Large Language Models</i>	
Dang Van Thin, Cuong Xuan Chu, Christian Graf, Tobias Kaminski and Trung-Kien Tran . .	1588
<i>Building Data-Driven Occupation Taxonomies: A Bottom-Up Multi-Stage Approach via Semantic Clustering and Multi-Agent Collaboration</i>	
Nan Li, Bo Kang and Tijl De Bie	1596
<i>AutoPenBench: A Vulnerability Testing Benchmark for Generative Agents</i>	
Luca Gioacchini, Alexander Delsanto, Idilio Drago, Marco Mellia, Giuseppe Siracusano and Roberto Bifulco	1615
<i>Enabling Self-Improving Agents to Learn at Test Time With Human-In-The-Loop Guidance</i>	
Yufei He, Ruoyu Li, Alex Chen, Yue Liu, Yulin Chen, Yuan Sui, Cheng Chen, Yi Zhu, Luca Luo, Frank Yang and Bryan Hooi	1625
<i>Encouraging Good Processes Without the Need for Good Answers: Reinforcement Learning for LLM Agent Planning</i>	
Zhiwei Li, Yong Hu and Wenqing Wang	1654
<i>Experience Report: Implementing Machine Translation in a Regulated Industry</i>	
Marco Zocca, Per Fallgren and David Buffoni	1667
<i>Multi-Task Pre-Finetuning of Lightweight Transformer Encoders for Text Classification and NER</i>	
Junyi Zhu, Savas Ozkan, Andrea Maracani, Sinan Mutlu, Cho Jung Min and Mete Ozay . . .	1674
<i>Scaling Down, Serving Fast: Compressing and Deploying Efficient LLMs for Recommendation Systems</i>	
Kayhan Behdin, Ata Fatahibaarzi, Qingquan Song, Yun Dai, Aman Gupta, Zhipeng Wang, Hejian Sang, Shao Tang, Gregory Dexter, Sirou Zhu, Siyu Zhu, Tejas Dharamsi, Vignesh Kothapalli, Zhoutong Fu, Yihan Cao, Pin-Lun Hsu, Fedor Borisjuk, Natesh S. Pillai, Luke Simon and Rahul Mazumder	1687
<i>Group, Embed and Reason: A Hybrid LLM and Embedding Framework for Semantic Attribute Alignment</i>	
Shramona Chakraborty, Shashank Mujumdar, Nitin Gupta, Sameep Mehta, Ronen Kat, Itay Etelis, Mohamed Mahameed, Itai Guez and Rachel Tzoref-Brill	1703

<i>STREQ: Selective Tiered Routing for Effective and Affordable Contact Center Quality Assurance</i>	Prajwal Sood, Rajdeep Agrawal, Mayank Sati, Digvijay Anil Ingle and Cijo George	1711
<i>Divide, Link, and Conquer: Recall-oriented Schema Linking for NL-to-SQL via Question Decomposition</i>	Kiran Pradeep, Kirushikesh DB, Nishtha Madaan, Sameep Mehta and Pushpak Bhattacharyya	1727
<i>Declarative Techniques for NL Queries over Heterogeneous Data</i>	Elham Khabiri, Jeffrey O. Kephart, Fenno F. Heath Iii, Srideepika Jayaraman, Yingjie Li, Fateh A. Tipu, Dhruv Shah, Achille Fokoue and Anu Bhamidipaty	1744
<i>Taxonomy of Comprehensive Safety for Clinical Agents</i>	Jean Seo, Hyunkyung Lee, Gibaeg Kim, Wooseok Han, Jaehyo Yoo, Seungseop Lim, Kihun Shin and Eunho Yang	1762
<i>Dr. Copilot: A Multi-Agent Prompt Optimized Assistant for Improving Patient-Doctor Communication in Romanian</i>	Andrei Niculae, Adrian Cosma, Cosmin Dumitrache and Emilian Radoi	1780
<i>Data-Efficient Automatic Prompt Optimization for Memory-Enhanced Conversational Agents</i>	Ervine Zheng, Yikuan Li, Geoffrey Jay Tso and Jilong Kuang	1793
<i>CLARITY: Clinical Assistant for Routing, Inference, and Triage</i>	Vladimir Shaposhnikov, Alexandr Nesterov, Iliia Kopanichuk, Ivan Bakulin, Zhelvakov Egor, Ruslan Abramov, Tsapieva Ekaterina Olegovna, Iaroslav Radionovich Bespalov, Dmitry V. Dylov and Ivan Oseledets	1805
<i>HalluDetect: Detecting, Mitigating, and Benchmarking Hallucinations in Conversational Systems in the Legal Domain</i>	Spandan Anaokar, Shrey Ganatra, Swapnil Bhattacharyya, Harshvivek Kashid, Shruthi N Nair, Reshma Sekhar, Siddharth Manohar, Rahul Hemrajani and Pushpak Bhattacharyya	1822
<i>How Accurate Are LLMs at Multi-Question Answering on Conversational Transcripts?</i>	Xiliang Zhu, Shi Zong and David Rossouw	1848
<i>AI Knowledge Assist: An Automated Approach for the Creation of Knowledge Bases for Conversational AI Agents</i>	Md Tahmid Rahman Laskar, Julien Bouvier Tremblay, Xue-Yong Fu, Cheng Chen and Shashi Bhushan TN	1856
<i>DACIP-RC: Domain Adaptive Continual Instruction Pre-Training via Reading Comprehension on Business Conversations</i>	Elena Khasanova, Harsh Saini, Md Tahmid Rahman Laskar, Xue-Yong Fu, Cheng Chen and Shashi Bhushan TN	1867
<i>Analysis of Automated Document Relevance Annotation for Information Retrieval in Oil and Gas Industry</i>	João Vitor Mariano Correia, Murilo Missano Bell, João Vitor Robiatti Amorim, Jonas Queiroz, Daniel Pedronette, Ivan Rizzo Guilherme and Felipe Lima de Oliveira	1878
<i>Mind the Query: A Benchmark Dataset towards Text2Cypher Task</i>	Vashu Chauhan, Shobhit Raj, Shashank Mujumdar, Avirup Saha and Anannay Jain	1890
<i>Deploying Tiny LLM Judges for Real-World Evaluation of Chart Models: Lessons Learned and Best Practices</i>		

Md Tahmid Rahman Laskar, Mohammed Saidul Islam, Ridwan Mahbub, Mizanur Rahman, Amran Bhuiyan, Israt Jahan, Mir Tafseer Nayeem, Shafiq Joty, Enamul Hoque and Jimmy Huang . . .	1906
<i>Agent-in-the-Loop: A Data Flywheel for Continuous Improvement in LLM-based Customer Support</i> Cen Zhao, Tiantian Zhang, Hanchen Su, Yufeng Zhang, Shaowei Su, Mingzhi Xu, Yu Liu, Wei Han, Jeremy Werner, Claire Na Cheng and Yashar Mehdad	1919
<i>Beyond Pointwise Scores: Decomposed Criteria-Based Evaluation of LLM Responses</i> Fangyi Yu, Nabeel Seedat, Drahomira Herrmannova, Frank Schilder and Jonathan Richard Schwarz	1931
<i>Scalable and Cost Effective High-Cardinality Classification with LLMs via Multi-View Label Representations and Retrieval Augmentation</i> Anup Pattnaik, Sasanka Vutla, Hamvir Dev, Jeevesh Nandan and Cijo George	1955
<i>How to Fine-Tune Safely on a Budget: Model Adaptation Using Minimal Resources</i> Anh C. Pham, Mihir Thalanki, Michael Sun, Aditya Chaloo, Ankita Gupta, Tian Xia, Aditya Mate, Ehi Nosakhare and Soundararajan Srinivasan	1970
<i>Zero-knowledge LLM hallucination detection and mitigation through fine-grained cross-model consistency</i> Aman Goel, Daniel Schwartz and Yanjun Qi	1982
<i>Incremental Summarization for Customer Support via Progressive Note-Taking and Agent Feedback</i> Yisha Wu, Cen Zhao, Yuanpei Cao, Xiaoqing Xu, Yashar Mehdad, Mindy Ji and Claire Na Cheng	2000
<i>LLMInit: A Free Lunch from Large Language Models for Selective Initialization of Recommendation</i> Weizhi Zhang, Liangwei Yang, Woosong Yang, Henry Peng Zou, Yuqing Liu, Ke Xu, Sourav Medya and Philip S. Yu	2016
<i>LLM Agents Implement an NLG System from Scratch: Building Interpretable Rule-Based RDF-to-Text Generators</i> Mateusz Lango and Ondrej Dusek	2025
<i>Leveraging LLMs to Streamline the Review of Public Funding Applications</i> João DS Marques, Andre Vicente Duarte, André Mendes Marques de Carvalho, Gil Rocha, Bruno Martins and Arlindo L. Oliveira	2041
<i>AMAS: Adaptively Determining Communication Topology for LLM-based Multi-agent System</i> Hui Yi Leong, Yuheng Li, Yuqing Wu, Wenwen Ouyang, Wei Zhu, Jiechao Gao and Wei Han	2061
<i>ColMate: Contrastive Late Interaction and Masked Text for Multimodal Document Retrieval</i> Ahmed Masry, Megh Thakkar, Patrice Bechard, Sathwik Tejaswi Madhusudhan, Rabiul Awal, Shambhavi Mishra, Akshay Kalkunte Suresh, Srivatsava Daruru, Enamul Hoque, Spandana Gella, Torsten Scholak and Sai Rajeswar	2071
<i>Confidence-Aware Reasoning: Optimizing Self-Guided Thinking Trajectories in Large Reasoning Models</i> Jiaxin Zhang	2081
<i>Multi-Value-Product Retrieval-Augmented Generation for Industrial Product Attribute Value Identification</i> Huike Zou, Haiyang Yang, Yindu Su, Chen Li Yu, Qinye Xie, Chengbao Lian, Qingheng Zhang, Shuguang Han, Fei Huang and Jufeng Chen	2096

<i>AttributeForge: An Agentic LLM Framework for Automated Product Schema Modeling</i>	
Yunhan Huang, Klevis Ramo, Andrea Iovine, Melvin Monteiro, Sedat Gokalp, Arjun Bakshi, Hasan Turalic, Arsh Kumar, Jona Neumeier, Ripley Yates, Rejaul Monir, Simon Hartmann, Tushar Manglik and Mohamed Yakout	2106
<i>VestaBench: An Embodied Benchmark for Safe Long-Horizon Planning Under Multi-Constraint and Adversarial Settings</i>	
Tanmana Sadhu, Yanan Chen and Ali Pesaranhader	2122
<i>Advancing E-commerce Merchants Telemarketing with Synthetic Data-Driven LLMs</i>	
Qi Gou, Zehua Xia, Li Juan, Qingyang Zhao and Wenjing Yang	2146
<i>Depression Detection on Social Media with Large Language Models</i>	
Xiaochong Lan, Zhiguang Han, Yiming Cheng, Li Sheng, Jie Feng, Chen Gao and Yong Li	2155
<i>BullyBench: Youth & Experts-in-the-loop Framework for Intrinsic and Extrinsic Cyberbullying NLP Benchmarking</i>	
Kanishk Verma, Sri Balaaji, Joachim Wagner, Arefeh Kazemi, Darragh Mccashin, isobel Walsh@dcu, Sayani Basak, Sinan Ascı, Yelena Cherkasova, Alexandros Poulis, James Ohiggins Norman, Rebecca Umbach Umbach, Tijana Milosevic and Brian Davis	2172
<i>Tagging-Augmented Generation: Assisting Language Models in Finding Intricate Knowledge In Long Contexts</i>	
Anwesan Pal, Karen Hovsepian, Tinghao Guo, Mengnan Zhao, Somendra Tripathi, Nikos Kanakaris, George Mihaila and Sumit Nigam	2209
<i>DispatchQA: A Benchmark for Small Function Calling Language Models in E-Commerce Applications</i>	
Joachim Daiber, Victor Maricato, Ayan Sinha and Andrew Rabinovich	2221
<i>Generalized Embedding Models for Industry 4.0 Applications</i>	
Christodoulos Constantinides, Shuxin Lin and Dhaval C Patel	2234
<i>ECHO-LLaMA: Efficient Caching for High-Performance LLaMA Training</i>	
Maryam Dialameh, Rezaul Karim, Hossein Rajabzadeh, Omar Mohamed Awad, Boxing Chen, Hyock Ju Kwon, Walid Ahmed and Yang Liu	2252
<i>Generating Spatial Knowledge Graphs from Automotive Diagrams for Question Answering</i>	
Steve Bakos, Chen Xing, Heidar Davoudi, Aijun An and Ron DiCarlantonio	2270
<i>Enhancing Persuasive Dialogue Agents by Synthesizing Cross-Disciplinary Communication Strategies</i>	
Shinnosuke Nozue, Yuto Nakano, Yotaro Watanabe, Meguru Takasaki, Shoji Moriya, Reina Akama and Jun Suzuki	2287
<i>BIOPSY - Biomarkers In Oncology: Pipeline for Structured Yielding</i>	
Sanya A. Chetwani and Jaseem Mahmmdda	2313
<i>DAST: Difficulty-Adaptive Slow-Thinking for Large Reasoning Models</i>	
Yi Shen, Jian Zhang, Jieyun Huang, Shuming Shi, Wenjing Zhang, Jiangze Yan, Ning Wang, Kai Wang, Zhaoxiang Liu and Shiguo Lian	2322
<i>pEBR: A Probabilistic Approach to Embedding Based Retrieval</i>	
Han Zhang, Yunjiang Jiang, Mingming Li, Haowei Yuan, Yiming Qiu and Wen-Yun Yang .	2332
<i>Finding Diamonds in Conversation Haystacks: A Benchmark for Conversational Data Retrieval</i>	
Yohan Lee, Yongwoo Song and Sangyeop Kim	2343

<i><SYNTACT>: Structuring Your Natural Language SOPs into Tailored Ambiguity-Resolved Code Templates</i>	
Sachin Kumar Giroh, Pushpendu Ghosh, Aryan Jain, Harshal Giridhari Paunikar, Aditi Rastogi, Promod Yenigalla and Anish Nediyanath	2367
<i>Recover-LoRA: Data-Free Accuracy Recovery of Degraded Language Models via Low-Rank Adaptation</i>	
Devleena Das, Rajeev Patwari and Ashish Sirasao	2377
<i>ixi-GEN: Efficient Industrial sLLMs through Domain Adaptive Continual Pretraining</i>	
Seonwu Kim, Yohan Na, Kihun Kim, Hanhee Cho, Geun Lim, Mintae Kim, Seongik Park, Ki Hyun Kim, Youngsub Han and Byoung-Ki Jeon	2387
<i>GRAFT: A Graph-based Flow-aware Agentic Framework for Document-level Machine Translation</i>	
Himanshu Dutta, Sunny Manchanda, Prakhar Bapat, Meva Ram Gurjar and Pushpak Bhattacharyya	2405
<i>Recon, Answer, Verify: Agents in Search of Truth</i>	
Satyam Shukla, Himanshu Dutta and Pushpak Bhattacharyya	2429
<i>T-VEC: A Telecom-Specific Vectorization Model with Enhanced Semantic Understanding via Deep Triplet Loss Fine-Tuning</i>	
Vignesh Ethiraj, Ashwath D, Sidhanth Menon, Divya Vijay and Vidhyakshaya Kannan	2449
<i>PlanGPT-VL: Enhancing Urban Planning with Domain-Specific Vision-Language Models</i>	
He Zhu, Junyou Su, Minxin Chen, Wen Wang, Yijie Deng, Guanhua Chen and Wenjia Zhang	2461
<i>IPR: Intelligent Prompt Routing with User-Controlled Quality-Cost Trade-offs</i>	
Aosong Feng, Balasubramaniam Srinivasan, Yun Zhou, Zhichao Xu, Kang Zhou, Sheng Guan, Yueyan Chen, Xian Wu, Ninad Kulkarni, Yi Zhang, Zhengyuan Shen, Dmitriy Bepalov, Soumya Smruti Mishra, Yifei Teng, Darren Yow-Bang Wang, Haibo Ding and Lin Lee Cheong	2484
<i>Semantic Agreement Enables Efficient Open-Ended LLM Cascades</i>	
Duncan Soiffer, Steven Kolawole and Virginia Smith	2499
<i>Lost in Pronunciation: Detecting Chinese Offensive Language Disguised by Phonetic Cloaking Replacement</i>	
Haotan Guo, Jianfei He, Jiayuan Ma, Hongbin Na, Zimu Wang, Haiyang Zhang, Qi Chen, Wei Wang, Zijing Shi, Tao Shen and Ling Chen	2538
<i>Distilling Cross-Modal Knowledge into Domain-Specific Retrievers for Enhanced Industrial Document Understanding</i>	
Jinhyeong Lim, Jeongwan Shin, Seeun Lee, Seongdeok Kim, Jounghsu Choi, Jongbae Kim, Chun Hwan Jung and Youjin Kang	2551
<i>Don't Forget the Base Retriever! A Low-Resource Graph-based Retriever for Multi-hop Question Answering</i>	
Andre Melo, Enting Chen, Pavlos Vougiouklis, Chenxin Diao, Shriram Piramanayagam, Ruofei Lai and Jeff Z. Pan	2564
<i>Beyond Dynamic Quantization: An Efficient Static Hierarchical Mix-precision Framework for Near-Lossless LLM Compression</i>	
Yi Zhang, Kai Zhang, Zheyang Li, Wenming Tan, Ye Ren and Jilin Hu	2573
<i>STACKFEED: Structured Textual Actor-Critic Knowledge base editing with FEEDback</i>	
Shashank Kirtania, Naman Gupta, Priyanshu Gupta, Sumit Gulwani, Arun Iyer, Suresh Parthasarathy Iyengar, Arjun Radhakrishna, Sriram K. Rajamani and Gustavo Soares	2588

<i>JaCorpTrack: Corporate History Event Extraction for Tracking Organizational Changes</i> Yuya Sawada, Hiroki Ouchi, Yuichiro Yasui, Hiroki Teranishi, Yuji Matsumoto, Taro Watanabe and Masayuki Ishii	2607
<i>CTR-Guided Generative Query Suggestion in Conversational Search</i> Erxue Min, Hsiu-Yuan Huang, Xihong Yang, Min Yang, Xin Jia, Yunfang Wu, Hengyi Cai, Jun- feng Wang, Shuaiqiang Wang and Dawei Yin	2624
<i>LATTE: Learning Aligned Transactions and Textual Embeddings for Bank Clients</i> Egor Fadeev, Dzhambulat Mollaev, Aleksei Shestov, Dima Korolev, Omar Zoloev, Ivan A Kireev, Andrey Savchenko and Maksim Makarenko	2635
<i>RedOne: Revealing Domain-specific LLM Post-Training in Social Networking Services</i> Fei Zhao, Chonggang Lu, Wangyue, Zheyong Xie, Ziyang Liu, Haofu Qian, Jianzhao Huang, Fangcheng Shi, Zijie Meng, Hongcheng Guo, Mingqian He, Xinze Lyu, Zheyu Ye, Weiting Liu, Boyang Wang and Shaosheng Cao	2648
<i>High-Quality Medical Dialogue Synthesis for Improving EMR Generation</i> Chengze Ge, Yu Xu, Qi Shao and Shengping Liu	2675
<i>ZI: Efficient Test-time Scaling with Code</i> Zhaojian Yu, Yinghao Wu, Yilun Zhao, Arman Cohan and Xiao-Ping Zhang	2688
<i>Quality Assessment of Tabular Data using Large Language Models and Code Generation</i> Ashlesha Akella, Akshar Kaul, Krishnasuri Narayanam and Sameep Mehta	2713
<i>PARSE: LLM Driven Schema Optimization for Reliable Entity Extraction</i> Anubhav Shrimal, Aryan Jain, Soumyajit Chowdhury and Promod Yenigalla	2749
<i>From Long Videos to Engaging Clips: A Human-Inspired Video Editing Framework with Multimodal Narrative Understanding</i> Xiangfeng Wang, Xiao Li, Yadong Wei, Songxueyu, Yang Song, Xiaoxiaoqiang, Fangrui Zeng, Zaiyi Chen, Liuliu, Gu Xu and Tong Xu	2764
<i>Efficiency-Effectiveness Reranking FLOPs for LLM-based Rerankers</i> Zhiyuan Peng, Ting-Ruen Wei, Tingyu Song and Yilun Zhao	2782
<i>GEAR: A Scalable and Interpretable Evaluation Framework for RAG-Based Car Assistant Systems</i> Niloufar Beyranvand, Hamidreza Dastmalchi, Aijun An, Heidar Davoudi, Winston Chan and Ron DiCarlantonio	2792
<i>FQ-Eval: Building Evaluation Dataset for User-centered Follow-up Question Generation</i> Sanghyun Seo, Bumsoo Kang, Dahm Lee, Jaeheon Kim, Joongbo Shin, Eui Soon Kim and Ki- jeong Jeon	2811
<i>Evaluating AI for Finance: Is AI Credible at Assessing Investment Risk Appetite?</i> Divij Chawla, Ashita Bhutada, Duc Anh Do, Abhinav Raghunathan, Vinod SP, Cathy Guo, Dar Win Liew, Prannaya Gupta, Rishabh Bhardwaj, Rajat Bhardwaj and Soujanya Poria	2828
<i>CAPSTONE: Composable Attribute-Prompted Scene Translation for Zero-Shot Vision–Language Rea- soning</i> Md. Ismail Hossain, Shahriyar Zaman Ridoy, Moshiur Farazi, Nabeel Mohammed and Shafin Rahman	2840
<i>Building Resource-Constrained Language Agents: A Korean Case Study on Chemical Toxicity Informa- tion</i> Hojun Cho, Donghu Kim, Soyoun Yang, Chan Lee, Hunjoo Lee and Jaegul Choo	2852

AutoDSPy: Automating Modular Prompt Design with Reinforcement Learning for Small and Large Language Models

Nafew Azim, Abrar Ur Alam, Hasan Bin Omar, Abdullah Mohammad Muntasir Adnan Jami, Jawad Ibn Ahad, Muhammad Rafsan Kabir, Md. Ismail Hossain, Fuad Rahman, Mohammad Ruhul Amin, Shafin Rahman and Nabeel Mohammed2881