

EMNLP 2024

**The 2024 Conference on Empirical Methods in Natural  
Language Processing**

**Proceedings of System Demonstrations**

November 12-16, 2024

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

## Diamond



CITADEL | CITADEL Securities

## Platinum



Megagon Labs



## Gold



## Silver



## Bronze

JPMorganChase



## D&I Champion



©2024 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](mailto:acl@aclweb.org)

ISBN 979-8-89176-167-4

## Introduction

Welcome to the proceedings of the system demonstration track of the 2024 Conference on Empirical Methods in Natural Language Processing on November 12th – November 16th, 2024. For the EMNLP 2024 system demonstration track, we received 153 submissions, of which 52 were selected for inclusion in the program (acceptance rate of 34%). This year, we also recruited area chairs in the system demonstration track. We would like to thank the area chairs and all members of the program committee for their help in reviewing the submissions. Lastly, we thank the many authors who submitted their work to the demonstrations track. This year, the EMNLP conference is a hybrid event. The demonstration papers will be presented through pre-recorded talks and in presence during the poster sessions.

Manling Li, Delia Irazú Hernández Farías, Tom Hope  
EMNLP 2024 System Demonstration Chairs

## Program Committee

### Chairs

Manling Li, Northwestern University  
Delia Irazu Hernandez Farias, INAOE  
Tom Hope, AI2, HUJI

### Area Chairs

Carl Edwards, University of Illinois at Urbana-Champaign  
Chi Han, University of Illinois, Urbana Champaign  
Denghui Zhang, Stevens Institute of Technology  
Emma Strubell, Allen Institute for Artificial Intelligence  
Hiba Arnaout, Technische Universität Darmstadt  
Hua Wei, Arizona State University  
Huajie Shao, College of William and Mary  
Ji-Ung Lee, Saarland University, Universität des Saarlandes  
Junxian He, Hong Kong University of Science and Technology  
Kuan-Hao Huang, Texas A&M University  
Lianhui Qin, University of California, San Diego  
Long Bai, Institute of Computing Technology, Chinese Academy of Sciences  
Qingyun Wang, University of Illinois, Urbana Champaign  
Shizhe Diao, Hong Kong University of Science and Technology  
Simra Shahid, University of Virginia, Charlottesville  
Tirthankar Ghosal, Oak Ridge National Laboratory  
Xinya Du, University of Texas at Dallas  
Yaojie Lu, Institute of Software, Chinese Academy of Sciences  
Yi Fung, University of Illinois at Urbana-Champaign  
Yu Su, Ohio State University  
Zejiang Shen, Massachusetts Institute of Technology  
Zhaozhuo Xu, Stevens Institute of Technology  
Zixuan Li, Institute of Computing Technology, Chinese Academy of Sciences

### Program Committee

Aakanksha Naik, Allen Institute for Artificial Intelligence  
Aarne Talman, University of Helsinki  
Adithya Pratapa, Carnegie Mellon University  
Alay Dilipbhai Shah, Tensoropera AI  
Alda Mari, CNRS  
Alessandra Teresa Cignarella, Ghent University  
Alexander M Rush, Cornell University  
Alexander Spangher, University of Southern California  
Amanda Bertsch, Carnegie Mellon University  
Amita Misra, Amazon  
Anirudh Ajith, Princeton University  
Ante Wang, Xiamen University  
Arda Tezcan, Universiteit Gent

Arie Cattan, Bar Ilan University  
Aryan Singhal, Carnegie Mellon University  
Bonaventura Coppola, SAP Security Research  
Boxi Cao, Institute of Software, Chinese Academy of Sciences  
Bram Vanroy, Instituut voor de Nederlandse Taal  
Carl Edwards, University of Illinois at Urbana-Champaign  
Chan-Wei Hu, Texas A&M University - College Station  
Charles Yu, University of Illinois at Urbana-Champaign  
Chen Shani, Stanford University  
Chen Zhang, Peking University  
Cheng Qian, University of Illinois at Urbana-Champaign  
Chengjie Sun, Harbin Institute of Technology  
Chenyang Zhao, UCLA Computer Science Department, University of California, Los Angeles  
Chujie Zheng, Tsinghua University  
Chung-Chi Chen, AIST, National Institute of Advanced Industrial Science and Technology  
Dayu Hu, National University of Defense Technology  
Duy Le, Rice University  
Emily Herron, Oak Ridge National Laboratory  
Enes Yavuz Ugan, Karlsruher Institut für Technologie  
Eugenio Martínez-Cámara, Universidad de Jaén  
Fabian Retkowski, Karlsruher Institut für Technologie  
Fan Yang, Ohio State University, Columbus  
Farah Benamara, Institut de recherche en informatique de toulouse  
Georgeta Bordea, Université de La Rochelle  
Gosuddin Kamaruddin Siddiqi, Microsoft  
Gábor Berend, University of Szeged  
Hao Yu, National University of Defense Technology  
Haoran Ranran Zhang, Pennsylvania State University  
Hiba Arnaout, Technische Universität Darmstadt  
Hillel Taub-Tabib, Allen Institute for Artificial Intelligence  
Hongyi Liu, Rice University  
Huy V. Nguyen, Amazon  
Hyewon Jeong, Massachusetts Institute of Technology  
I-Hung Hsu, Google  
Isadora White, University of California, San Diego  
Jeonghwan Kim, University of Illinois at Urbana-Champaign  
Jeremiah Milbauer, Carnegie Mellon University  
Jian Zheng, Amazon AGI  
Jiangming Liu, Yunnan University  
Jianyi Zhang, Duke University  
Jiateng Liu, Department of Computer Science  
Jiaxin Pei, Stanford University  
Jiazhan Feng, Peking University  
John Philip McCrae, National University of Ireland Galway  
Junpeng Liu, The Chinese University of Hong Kong  
Jyotika Singh, Placemakr  
Ke Liang, National University of Defense Technology  
Kai Sun, Meta  
Kangrui Wang, Northwestern University  
Ke Yang, Department of Computer Science  
Kehai Chen, Harbin Institute of Technology (Shenzhen)

Lei Shu, Google  
Lifeng Han, Leiden University, Leiden University  
Liliang Ren, Microsoft GenAI  
Linfan Zhang, University of California, Los Angeles  
LingXi Zhang, Rice University  
Lingyuan Meng, National University of Defense Technology  
Lixin Su, Baidu  
Long Bai, Institute of Computing Technology, Chinese Academy of Sciences  
Longchao Da, Arizona State University  
Luigi Di Caro, University of Turin, Italy  
Maarit Koponen, University of Eastern Finland  
Mamoru Komachi, Hitotsubashi University  
Marco Turchi, Zoom  
Marek Suppa, Comenius University in Bratislava  
Maria Mahbub, Oak Ridge National Laboratory  
Mario Ezra Aragon, Universidad de Santiago de Compostela  
Marissa Radensky, University of Washington  
Meilong Xu, State University of New York at Stony Brook  
Meishan Zhang, Harbin Institute of Technology (Shenzhen), China  
Meng Liu, National University of Defense Technology  
Miao Su, Institute of Computing Technology, Chinese Academy of Sciences  
Mingxu Tao, Peking University  
Monica Agrawal, Duke University  
Nathan Lambert, Allen Institute for Artificial Intelligence  
Nir Mazor, Hebrew University of Jerusalem  
Pegah Ahadian, Kent State University  
Pengfei Yu, Boson AI  
Pin Ni, University College London, University of London  
Qingkai Zeng, University of Notre Dame  
Qingyun Wang, University of Illinois, Urbana Champaign  
Qiusi Zhan, University of Illinois Urbana-Champaign  
Quzhe Huang, Peking University  
Raymond Fok, University of Washington  
Renxiang Guan, National University of Defense Technology  
Revanth Gangi Reddy, University of Illinois at Urbana-Champaign  
Rik Koncel-Kedziorski, Apple  
Samar Haider, University of Pennsylvania  
Sanyuan Chen, Microsoft  
Sha Li, University of Illinois, Urbana Champaign  
Shangqing Tu, Tsinghua University  
Shanshan Han, University of California, Irvine  
Shaochen Zhong, Rice University  
Shiguang Guo, University of the Chinese Academy of Sciences  
Shuming Shi, Tencent AI Lab  
Shuo Xing, Texas A&M University - College Station  
Sireesh Gururaja, School of Computer Science, Carnegie Mellon University  
Siru Ouyang, University of Illinois Urbana-Champaign Champaign  
Sizhe Zhou, University of Illinois Urbana-Champaign  
Steven Kolawole, Carnegie Mellon University  
Tal August, Allen Institute for Artificial Intelligence  
Tanay Dixit, University of illinois urbana champaign



Tianlin Zhang, University of Manchester  
Tianrui Liu, Microsoft  
Tiantian Zhu, Harbin Institute of Technology (Shenzhen)  
Tianyu Liu, Alibaba Group  
Tiejin Chen, Arizona State University  
Viet Dac Lai, Adobe Systems  
Voula Giouli, Aristotle University of Thessaloniki  
Wei Wang, Apple AI/ML  
Wei Xu, Georgia Institute of Technology  
Weicheng Ren, Chinese Academy of Sciences  
Weijiang Li, Harvard Medical School, Harvard University  
Wen Zhang, Zhejiang University  
William Merrill, New York University  
Xianrui Zhong, University of Illinois at Urbana-Champaign  
Xiao Liu, Microsoft Research Asia  
Xiaochang Li, William and Mary  
Xiaomeng Jin, University of Illinois at Urbana-Champaign  
Xiaozhi Wang, Department of Computer Science and Technology, Tsinghua University  
Xin Xu, The Hong Kong University of Science and Technology  
Xinchi Chen, Amazon  
Xintong Li, Apple  
Xinyu Lu, Institute of Software, Chinese Academy of Sciences  
Xueqing Wu, UCLA Computer Science Department, University of California, Los Angeles  
Xueying Jia, Carnegie Mellon University  
Yangyi Chen, Department of Computer Science, University of Illinois at Urbana-Champaign  
Yanjiang Liu, University of the Chinese Academy of Sciences  
Yanzhou Pan, Google  
Yaojie Lu, Institute of Software, Chinese Academy of Sciences  
Yekun Chai, Baidu  
Yevgeniy Puzikov, Zalando SE  
Yi Cai, South China University of Technology  
Yide Ran, Stevens Institute of Technology  
Yige Yuan, Institute of Computing Technology, Chinese Academy of Sciences  
Yiheng Shu, The Ohio State University  
Yiwei Wang, University of California, Merced  
Yoshihide Kato, Nagoya University  
Youyang Ng, Kioxia Corporation  
Yuan Sun, Sichuan University  
Yucheng Shi, University of Georgia  
Yudong Tao, Facebook  
Yue Huang, University of Notre Dame  
Yueqian Lin, Duke University  
Yufeng Du, University of Illinois at Urbana-Champaign  
Yuheng Zha, University of California, San Diego  
Yuji Zhang, Dalian University of Technology  
Yung-Sung Chuang, Massachusetts Institute of Technology  
Yutong Hu, Peking University  
Yuxuan Lai, The Open University of China  
Yuxuan Lu, Northeastern University  
Yuxuan Wan, Department of Computer Science and Engineering, The Chinese University of Hong Kong

Zheng Liu, Northeastern University  
Zhenhailong Wang, University of Illinois Urbana-Champaign  
Zhenyu Wu, University of Notre Dame  
Zhenyu Zhang, University of Texas at Austin  
Zhibin Lan, Xiamen University  
Zhiqiu Lin, Carnegie Mellon University  
Zhoujun Cheng, UC San Diego  
Zhuo Chen, Zhejiang University  
Zi-Yi Dou, University of California, Los Angeles  
Zichao Li, Institute of Software, Chinese Academy of Sciences  
Zifeng Wang, Google  
Zihan Wang, Northwestern University, Northwestern University  
Zineng Tang, University of North Carolina, Chapel Hill  
Ziqian Luo, Oracle  
Zirui Wu, Peking University  
Zixuan Li, Institute of Computing Technology, Chinese Academy of Sciences  
Zongyu Wu, Pennsylvania State University

## Table of Contents

<i>FreeEval: A Modular Framework for Trustworthy and Efficient Evaluation of Large Language Models</i> Zhuohao Yu, Chang Gao, Wenjin Yao, Yidong Wang, Zhengran Zeng, Wei Ye, Jindong Wang, Yue Zhang and Shikun Zhang .....	1
<i>i-Code Studio: A Configurable and Composable Framework for Integrative AI</i> Yuwei Fang, Mahmoud Khademi, Chenguang Zhu, Ziyi Yang, Reid Pryzant, Yichong Xu, Yao Qian, Takuya Yoshioka, Lu Yuan, Michael Zeng and Xuedong Huang .....	14
<i>Evalverse: Unified and Accessible Library for Large Language Model Evaluation</i> Jihoo Kim, Wonho Song, Dahyun Kim, Yunsu Kim, Yungi Kim and Chanjun Park .....	25
<i>Medico: Towards Hallucination Detection and Correction with Multi-source Evidence Fusion</i> Xinping Zhao, Jindi Yu, Zhenyu Liu, Jifang Wang, Dongfang Li, Yibin Chen, Baotian Hu and Min Zhang .....	34
<i>OpenOmni: A Collaborative Open Source Tool for Building Future-Ready Multimodal Conversational Agents</i> Qiang Sun, Yuanyi Luo, Sirui Li, Wenxiao Zhang and Wei Liu .....	46
<i>Lighthouse: A User-Friendly Library for Reproducible Video Moment Retrieval and Highlight Detec- tion</i> Taichi Nishimura, Shota Nakada, Hokuto Munakata and Tatsuya Komatsu .....	53
<i>MarkLLM: An Open-Source Toolkit for LLM Watermarking</i> Leyi Pan, Aiwei Liu, Zhiwei He, Zitian Gao, Xuandong Zhao, Yijian LU, Binglin Zhou, Shuliang Liu, Xuming Hu, Lijie Wen, Irwin King and Philip S. Yu .....	61
<i>AUTOGEN STUDIO: A No-Code Developer Tool for Building and Debugging Multi-Agent Systems</i> Victor Dibia, Jingya Chen, Gagan Bansal, Suff Syed, Adam Fourney, Erkang Zhu, Chi Wang and Saleema Amershi .....	72
<i>TinyAgent: Function Calling at the Edge</i> Lutfi Eren Erdogan, Nicholas Lee, Siddharth Jha, Sehoon Kim, Ryan Tabrizi, Suhong Moon, Coleman Richard Charles Hooper, Gopala Anumanchipalli, Kurt Keutzer and Amir Gholami .....	80
<i>TruthReader: Towards Trustworthy Document Assistant Chatbot with Reliable Attribution</i> Dongfang Li, Xinshuo Hu, Zetian Sun, Baotian Hu, Shaolin Ye, Zifei Shan, Qian Chen and Min Zhang .....	89
<i>Commentator: A Code-mixed Multilingual Text Annotation Framework</i> Rajvee Sheth, Shubh Nisar, Heenaben Prajapati, Himanshu Beniwal and Mayank Singh .....	101
<i>Integrating INCEPTION into larger annotation processes</i> Richard Eckart De Castilho, Jan-Christoph Klie and Iryna Gurevych .....	110
<i>Arxiv Copilot: A Self-Evolving and Efficient LLM System for Personalized Academic Assistance</i> Guanyu Lin, Tao Feng, Pengrui Han, Ge Liu and Jiaxuan You .....	122
<i>TransAgents: Build Your Translation Company with Language Agents</i> Minghao Wu, Jiahao Xu and Longyue Wang .....	131
<i>Monitoring Hate Speech in Indonesia: An NLP-based Classification of Social Media Texts</i> Musa Izzanardi Wijanarko, Lucky Susanto, Prasetia Anugrah Pratama, Ika Karlina Idris, Traci Hong and Derry Tanti Wijaya .....	142

<i>CAVA: A Tool for Cultural Alignment Visualization &amp; Analysis</i> Nevan Giuliani, Cheng Charles Ma, Prakruthi Pradeep and Daphne Ippolito . . . . .	153
<i>ReDel: A Toolkit for LLM-Powered Recursive Multi-Agent Systems</i> Andrew Zhu, Liam Dugan and Chris Callison-Burch . . . . .	162
<i>BattleAgent: Multi-modal Dynamic Emulation on Historical Battles to Complement Historical Analysis</i> Shuhang Lin, Wenyue Hua, Lingyao Li, Che-Jui Chang, Lizhou Fan, Jianchao Ji, Hang Hua, Mingyu Jin, Jiebo Luo and Yongfeng Zhang . . . . .	172
<i>sign.mt: Real-Time Multilingual Sign Language Translation Application</i> Amit Moryossef . . . . .	182
<i>WebOlympus: An Open Platform for Web Agents on Live Websites</i> Boyuan Zheng, Boyu Gou, Scott Salisbury, Zheng Du, Huan Sun and Yu Su . . . . .	187
<i>TAIL: A Toolkit for Automatic and Realistic Long-Context Large Language Model Evaluation</i> Gefei Gu, Yilun Zhao, Ruoxi Ning, Yanan Zheng and Arman Cohan . . . . .	198
<i>OpenResearcher: Unleashing AI for Accelerated Scientific Research</i> Yuxiang Zheng, Shichao Sun, Lin Qiu, Dongyu Ru, Cheng Jiayang, Xuefeng Li, Jifan Lin, Binjie Wang, Yun Luo, Renjie Pan, Yang Xu, Qingkai Min, Zizhao Zhang, Yiwen Wang, Wenjie Li and Pengfei Liu . . . . .	209
<i>OpenFactCheck: A Unified Framework for Factuality Evaluation of LLMs</i> Hasan Iqbal, Yuxia Wang, Minghan Wang, Georgi Nenkov Georgiev, Jiahui Geng, Iryna Gurevych and Preslav Nakov . . . . .	219
<i>ULLME: A Unified Framework for Large Language Model Embeddings with Generation-Augmented Learning</i> Hieu Man, Nghia Trung Ngo, Franck Dernoncourt and Thien Huu Nguyen . . . . .	230
<i>To the Globe (TTG): Towards Language-Driven Guaranteed Travel Planning</i> Da JU, Song Jiang, Andrew Cohen, Aaron Foss, Sasha Mitts, Arman Zharmagambetov, Brandon Amos, Xian Li, Justine T Kao, Maryam Fazel-Zarandi and Yuandong Tian . . . . .	240
<i>MATSA: Multi-Agent Table Structure Attribution</i> Puneet Mathur, Alexa Siu, Nedim Lipka and Tong Sun . . . . .	250
<i>OpenT2T: An Open-Source Toolkit for Table-to-Text Generation</i> Haowei Zhang, Shengyun Si, Yilun Zhao, Lujing Xie, Zhijian Xu, Lyuhao Chen, Linyong Nan, Pengcheng Wang, Xiangru Tang and Arman Cohan . . . . .	259
<i>ChatHF: Collecting Rich Human Feedback from Real-time Conversations</i> Andrew Li, Zhenduo Wang, Ethan Mendes, Duong Minh Le, Wei Xu and Alan Ritter . . . . .	270
<i>KMatrix: A Flexible Heterogeneous Knowledge Enhancement Toolkit for Large Language Model</i> Shun Wu, Di Wu, Kun Luo, XueYou Zhang, Jun Zhao and Kang Liu . . . . .	280
<i>Xinference: Making Large Model Serving Easy</i> Weizheng Lu, Lingfeng Xiong, Feng Zhang, Xuye Qin and Yueguo Chen . . . . .	291
<i>RETAIN: Interactive Tool for Regression Testing Guided LLM Migration</i> Tanay Dixit, Daniel Lee, Sally Fang, Sai Sree Harsha, Anirudh Sureshan, Akash V Maharaj and Yunyao Li . . . . .	301
<i>ClaimLens: Automated, Explainable Fact-Checking on Voting Claims Using Frame-Semantics</i> Jacob Devasier, Rishabh Mediratta, Phuong Anh Le, David Huang and Chengkai Li . . . . .	311

<i>RAGViz: Diagnose and Visualize Retrieval-Augmented Generation</i> Tevin Wang, Jingyuan He and Chenyan Xiong .....	320
<i>PyMarian: Fast Neural Machine Translation and Evaluation in Python</i> Thamme Gowda, Roman Grundkiewicz, Elijah Rippeth, Matt Post and Marcin Junczys-Dowmunt	328
<i>LLM-DetectAIve: a Tool for Fine-Grained Machine-Generated Text Detection</i> Mervat Abassy, Kareem Elozeiri, Alexander Aziz, Minh Ngoc Ta, Raj Vardhan Tomar, Bimarsha Adhikari, Saad El Dine Ahmed, Yuxia Wang, Osama Mohammed Afzal, Zhuohan Xie, Jonibek Mansurov, Ekaterina Artemova, Vladislav Mikhailov, Rui Xing, Jiahui Geng, Hasan Iqbal, Zain Muhammad Mujahid, Tarek Mahmoud, Akim Tsvigun, Alham Fikri Aji, Artem Shelmanov, Nizar Habash, Iryna Gurevych and Preslav Nakov .....	336
<i>Translation Canvas: An Explainable Interface to Pinpoint and Analyze Translation Systems</i> Chinmay Dandekar, Wenda Xu, Xi Xu, Siqi Ouyang and Lei Li .....	344
<i>mbrs: A Library for Minimum Bayes Risk Decoding</i> Hiroyuki Deguchi, Yusuke Sakai, Hidetaka Kamigaito and Taro Watanabe .....	351
<i>Debug Smarter, Not Harder: AI Agents for Error Resolution in Computational Notebooks</i> Konstantin Grotov, Artem Borzilov, Maksim Krivobok, Timofey Bryksin and Yaroslav Zharov	363
<i>Schema-Guided Culture-Aware Complex Event Simulation with Multi-Agent Role-Play</i> Sha Li, Revanth Gangi Reddy, Khanh Duy Nguyen, Qingyun Wang, Yi Fung, Chi Han, Jiawei Han, Kartik Natarajan, Clare R. Voss and Heng Ji .....	372
<i>SparkRA: A Retrieval-Augmented Knowledge Service System Based on Spark Large Language Model</i> Dayong Wu, Jiaqi Li, Baoxin Wang, Honghong Zhao, Siyuan Xue, Yanjie Yang, Zhijun Chang, Rui Zhang, Li Qian, Bo Wang, Shijin Wang, Zhixiong Zhang and Guoping Hu .....	382
<i>Generative Dictionary: Improving Language Learner Understanding with Contextual Definitions</i> Kai-Wen Tuan, Hai-Lun Tu and Jason S. Chang .....	390
<i>WalledEval: A Comprehensive Safety Evaluation Toolkit for Large Language Models</i> Prannaya Gupta, Le Qi Yau, Hao Han Low, I-Shiang Lee, Hugo Maximus Lim, Yu Xin Teoh, Koh Jia Hng, Dar Win Liew, Rishabh Bhardwaj, Rajat Bhardwaj and Soujanya Poria .....	397
<i>RAGLAB: A Modular and Research-Oriented Unified Framework for Retrieval-Augmented Generation</i> Xuanwang Zhang, Yun-Ze Song, Yidong Wang, Shuyun Tang, Xinfeng Li, Zhengran Zeng, Zhen Wu, Wei Ye, Wenyuan Xu, Yue Zhang, Xinyu Dai, Shikun Zhang and Qingsong Wen .....	408
<i>AutoTrain: No-code training for state-of-the-art models</i> Abhishek Thakur .....	419
<i>Sailor: Open Language Models for South-East Asia</i> Longxu Dou, Qian Liu, Guangtao Zeng, Jia Guo, Jiahui Zhou, Xin Mao, Ziqi Jin, Wei Lu and Min Lin .....	424
<i>RepoAgent: An LLM-Powered Open-Source Framework for Repository-level Code Documentation Generation</i> Qinyu Luo, Yining Ye, Shihao Liang, Zhong Zhang, Yujia Qin, Yaxi Lu, Yesai Wu, Xin Cong, Yankai Lin, Yingli Zhang, Xiaoyin Che, Zhiyuan Liu and Maosong Sun .....	436

<i>DeepPavlov 1.0: Your Gateway to Advanced NLP Models Backed by Transformers and Transfer Learning</i>	
Maksim Savkin, Anastasia Voznyuk, Fedor Ignatov, Anna Korzanova, Dmitry Karpov, Alexander Popov and Vasily Kononov .....	465
<i>Kandinsky 3: Text-to-Image Synthesis for Multifunctional Generative Framework</i>	
Arkhipkin Vladimir, Viacheslav Vasilev, Andrei Filatov, Igor Pavlov, Julia Agafonova, Nikolai Gerasimenko, Anna Averchenkova, Evelina Mironova, Bukashkin Anton, Konstantin Kulikov, Andrey Kuznetsov and Denis Dimitrov .....	475
<i>MIMIR: A Customizable Agent Tuning Platform for Enhanced Scientific Applications</i>	
Xiangru Tang, Chunyuan Deng, Hanminwang Hanminwang, Haoran Wang, Yilun Zhao, Wenqi Shi, Yi Fung, Wangchunshu Zhou, Jiannan Cao, Heng Ji, Arman Cohan and Mark Gerstein .....	486
<i>WildVis: Open Source Visualizer for Million-Scale Chat Logs in the Wild</i>	
Yuntian Deng, Wenting Zhao, Jack Hessel, Xiang Ren, Claire Cardie and Yejin Choi .....	497
<i>Instruction-Driven Game Engine: A Poker Case Study</i>	
Hongqiu Wu, Xingyuan Liu, Yan Wang and Hai Zhao .....	507
<i>LM-Interview: An Easy-to-use Smart Interviewer System via Knowledge-guided Language Model Exploitation</i>	
Hanming Li, Jifan Yu, Ruimiao Li, Zhanxin Hao, Yan Xuan, Jiayi Yuan, Bin Xu, Juanzi Li and Zhiyuan Liu .....	520