

ACL 2026

**The 64th Annual Meeting of the Association for  
Computational Linguistics (ACL 2026)**

**System Demonstrations**

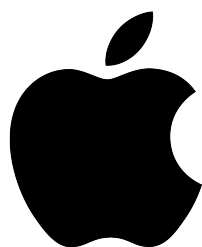
July 2-7, 2026

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

### Diamond



### Platinum



## Gold



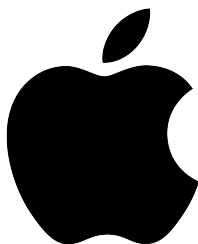
## Silver



## Bronze



## Academic Ally



©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](mailto:acl@aclweb.org)

ISBN 979-8-89176-392-0

## Introduction

Welcome to the proceedings of the System Demonstration Track of the 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026), held from July 2–7, 2026 in San Diego, California, United States.

The ACL 2026 System Demonstration Track provides a platform for papers describing system demonstrations, ranging from early research prototypes to mature, production-ready systems. Publicly available open-source or open-access systems are of special interest.

For the ACL 2026 System Demonstration Track, we received 227 submissions, of which 215 papers were valid with all required materials. Based on the submitted reviews, we accepted 85 papers, resulting in an acceptance rate of 37.45%.

From the accepted papers, we selected one paper to receive the Best Demo award: The olmOCR Project: Building Fully Open OCR using VLMs.

We sincerely thank the authors of all System Demonstration Track submissions as well as the reviewers and Area Chairs for their hard and dedicated work against tight deadlines. Our gratitude also extends to the General Chair and the Publication Chairs for their support throughout the compilation of this volume. Finally, we are grateful to all other ACL 2026 committees we collaborated with during all stages of conference organization and preparation.

Greg Durrett

Ping Jian

ACL 2026 Demonstration Chairs

# Organizing Committee

## General Chair

Philipp Koehn, Johns Hopkins University, USA

## Program Chairs

Maria Liakata, Queen Mary University of London, United Kingdom

Viviane P. Moreira, Institute of Informatics at UFRGS, Brazil

Jiajun Zhang, Chinese Academy of Sciences, China

David Jurgens, University of Michigan, USA

## Workshop Chairs

Loic Barrault, Meta, France

Yang Zhao, Chinese Academy of Sciences, China

## Tutorial Chairs

Kenton Murray, Johns Hopkins University, USA

Jacob Andreas, Massachusetts Institute of Technology, USA

## Demonstration Chairs

Greg Durrett, New York University, USA

Ping Jian, Beijing Institute of Technology, China

## Student Research Workshop Chairs

Santosh T.Y. S.S., Technical University of Munich and Amazon, Germany

Juan Diego Rodriguez, University of Texas at Austin, USA

Ona de Gibert, University of Helsinki, Finland

## Student Research Workshop Chairs - Faculty Advisors

Yves Scherrer, University of Oslo, Norway

Elias Stengel-Eskin, University of Texas at Austin, USA

Snigdha Chaturvedi, University of North Carolina, USA

## Publication Chairs

David Stap, NXAI

Haoran Xu, Microsoft, USA

Jan Niehues, Karlsruhe Institute of Technology, Germany

## Handbook Chairs

Jia Xu, Stevens Institute of Technology, USA

Ryan Cotterell, ETH Zürich, Switzerland

### **Sponsorship Chairs**

Baotian Hu, Harbin Institute of Technology, China  
Steve Richardson, Brigham Young University, USA

### **Diversity and Inclusion Chairs**

Ximena Gutierrez-Vasques, Universidad Nacional Autonoma de Mexico, Mexico  
Anietie Andy, Howard University, USA  
Gabrielle Gaudeau, University of Cambridge, United Kingdom  
Felermino Ali, Microsoft Research Africa, Kenya

### **Publicity Chairs**

Raj Dabre, Google, Australia  
Wei Xu, Georgia Tech, USA  
Flor Miriam Plaza-del-Arco, University of Leiden, Netherlands  
Debora Nozza, Bocconi University, Italy  
Yaping Zhang, Chinese Academy of Sciences, China

### **Website and Conference App Chairs**

Abulhair Saparov, Purdue University, USA  
Najoung Kim, Boston University, USA

### **Ethics Chairs**

Margaret Mitchell, Huggingface, USA  
Paco Guzman, Meta, USA

### **Student Volunteer Chairs**

Jingbo Shang, University of California at San Diego, USA  
Lianhui Qin, University of California at San Diego, USA  
Xuezhe Ma, University of Southern California, USA  
Muhao Chen, University of California at Davis, USA

### **Technical Open Review Chairs**

Shuoyang Ding, NVIDIA, USA  
Yi R. Fung, HKUST, Hong Kong, China

### **Virtual Infrastructure Chairs**

Gael Guibon, Université Sorbonne Paris Nord, France  
Parisa Kordjamshidi, Michigan State University, USA

### **Internal Communication Chairs**

Ming Jiang, Indiana University Indianapolis, USA  
Margot Mieskes, University of Applied Sciences, Darmstadt, Germany

### **Industry Track Chairs**

Georg Rehm, German Research Institute for Artificial Intelligence, Germany  
Yunyao Li, Adobe, USA  
Mei Tu, Samsung, China

### **Best Paper Chairs**

Antonios Anastasopoulos, George Mason University, USA  
Eunsol Choi, New York University, USA

### **Visa Chairs**

David Chiang, University of Notre Dame, USA  
Swabha Swayamdipta, University of Southern California, USA

### **Conflict Chair**

Isabelle Augenstein, University of Copenhagen, Denmark

# Program Committee

## Program Chairs

Greg Durrett, New York University  
Ping Jian, Beijing Institute of Technology

## Area Chairs

Emily Allaway, University of Edinburgh  
Milad Alshomary  
Yufeng Chen  
Caio Corro, Sorbonne Université  
Shizhe Diao, NVIDIA  
Xinya Du, University of Texas at Dallas  
Andre Freitas, Idiap Research Institute and University of Manchester  
Daniel Fried, Meta AI and Carnegie Mellon University  
Dan Goldwasser, Purdue University and Purdue University  
Vivek Gupta, Arizona State University  
Xianpei Han, Institute of Software, CAS  
Hyeju Jang, Indiana University  
Sachin Kumar, Ohio State University, Columbus  
Jonathan K. Kummerfeld, University of Sydney  
Shou-De Lin, National Taiwan University and National Taiwan University  
Lajanugen Logeswaran, LG AI Research  
Mounica Maddela, Bloomberg  
Kelly Marchisio, Cohere and Cohere  
Ashutosh Modi, IIT Kanpur  
Yixin Nie, Meta Platforms, Inc.  
Archiki Prasad  
Huajie Shao, College of William and Mary  
Nicholas Tomlin, Toyota Technological Institute at Chicago  
Hua Wei, Arizona State University and Arizona State University  
Xin Xin, Beijing Institute of Technology  
Zhaozhuo Xu, Workato and Stevens Institute of Technology  
Yang Zhang, International Business Machines  
Rui Zhang, Pennsylvania State University  
Zhisong Zhang, City University of Hong Kong  
Denghui Zhang, Stevens Institute of Technology  
Jieyu Zhao, University of Southern California

## Reviewers

Marek Šuppa

Amey Agrawal, Sydney Anuyah, Hiba Arnaout, Berk Atıl, Tal August

Fan Bai, Long Bai, Shweta Bhardwaj, Jing Bi, Georgeta Bordea, Eric Boxer, Nimet Beyza Bozdog

Boxi Cao, Danilo Carvalho, Yuanjun Chai, Jixuan Chen, Kehai Chen, Tiejin Chen, Xinchu Chen, Yubo Chen, Yuefei Chen, Yufeng Chen, Zhuoming Chen, Bonaventura Coppola

Longchao Da, Yichuan Deng, Zhiyang Deng, Luigi Di Caro, Wenxuan Ding, Michael Dominguez, Runpei Dong, Shihan Dou, Congrui Du

Mohamed Elgaar

Kuofeng Gao, Albert Ge, Debela Gemechu, Voula Giouli, Sebastian Gombert, Shanshan Gong, Renxiang Guan, Zhongbin Guo

Bin Han, Shanshan Han, Shirley Anugrah Hayati, Bingxiang He, Kaiyuan Hou, I-Hung Hsu, Chan-Wei Hu, Lanxiang Hu, Qi Yang Huang, Yuzhen Huang

Abhinav Jain, Jiabao Ji, Mohan Jiang, Han Jin

Ryo Kamoi, Wan Ju Kang, Yoshihide Kato, Patrik Kenfack, Jeonghwan Kim, Steven Kolawole, Varun Kotte, Satyapriya Krishna, Lun-Wei Ku, Nalin Kumar

Yuhang Lai, Yuxuan Lai, Tian Lan, Yueh Che Lee, Hengli Li, Ryan Li, Weijiang Li, Xiaochang Li, Yanshu Li, Yi-Chen Li, Yingya Li, Yixing Li, Zixuan Li, Yueqian Lin, Tianrui Liu, Xuefeng Liu, Yanjiang Liu, Yuhan Liu, Anamika Lochab, Dunjie Lu, Yaojie Lu, Ziqian Luo

Ming Ma, Alda Mari, John Philip McCrae, Jeremiah Milbauer, Amita Misra, Diganta Misra, Yash More

Deepak Nathani, Dianwen Ng, Youyang Ng, Dang Nguyen, Huy V. Nguyen, Pin Ni, Yuzhou Nie, Wanting Ning

Timothy Obiso

Soumyadeep Pal, ChaeHun Park, Core Francisco Park, Chau Minh Pham, Akshara Prabhakar

Cheng Qian

Chahat Raj, Kaushik Roy

Ali Riahi Samani, Jorge Sanchez, Alay Dilipbhai Shah, Yiran Jenny Shen, Zhennan Shen, Dongchan Shin, Gosuddin Kamaruddin Siddiqi, Jyotika Singh, Prasann Singhal, Yejin Son, Nathan Stringham, Si Sun

Sotaro Takeshita, Yuzhi Tang, Mingxu Tao, Yudong Tao, Runchu Tian

Julien Vanegue, Ivo Verhoeven

Guangya Wan, Jiong Xiao Wang, Junke Wang, Qingyun Wang, Shida Wang, Wei Wang, Zihan Wang, Franziska Weeber, Xilin Wei, Mengsong Wu, Qiucheng Wu, Zekun Wu, Zhenyu Wu

Shifeng Xie, Xin Xin, Shuo Xing, Hao Xu, Meilong Xu, Xin Xu, Zhenran Xu

Zhen Yang, Zonghao Ying, Hao Yu, Pengfei Yu

Qingkai Zeng, Yuheng Zha, Chen Zhang, Juzheng Zhang, Linfan Zhang, Lingxi Zhang, Tianlin Zhang, Tianyang Zhang, Yunxiang Zhang, Zizhuo Zhang, Bingchen Zhao, Jian Zheng, Yang Zhong, Guochen Zhou, Sizhe Zhou, Zhanke Zhou, Tiantian Zhu, Zhou Ziheng

## Table of Contents

<i>Interpreto: An Explainability Library for Transformers</i> Antonin Poché, Thomas Mullor, Gabriele Sarti, Frédéric Boissard, Corentin Friedrich, Charlotte Claye, Francois Hoofd, Raphael Bernas, Nicholas Asher, Celine Hudelot and Fanny Jourdan . . . . .	1
<i>Copyright Detective: A Forensic System to Evidence LLMs Flickering Copyright Leakage Risks</i> Guangwei Zhang, Jianing Zhu, Cheng Qian, Neil Zhenqiang Gong, Rada Mihalcea, Zhaozhuo Xu, Jingrui He, Jiaqi W. Ma, Chaowei Xiao, Bo Li, Ahmed Abbasi, Dongwon Lee, Heng Ji and Denghui Zhang . . . . .	14
<i>PROTEA: Offline Evaluation and Iterative Refinement for Multi-Agent LLM Workflows</i> Kazuki Kawamura, Satoshi Waki and Kei Tateno . . . . .	27
<i>SymbolicThought: Integrating Language Models and Symbolic Reasoning for Consistent and Interpretable Human Relationship Understanding</i> Runcong Zhao, Qinglin Zhu, Hainiu Xu, Bin Liang, Lin Gui and Yulan He . . . . .	36
<i>LiTS: A Modular Framework for LLM Tree Search</i> Xinzhe Li and Yaguang Tao . . . . .	47
<i>AutoFigure-Edit: Generating Editable Scientific Illustrations via Reference-Guided Styling</i> Zhen Lin, Qiujie Xie, Minjun Zhu, Shichen Li, QiYao Sun, Enhao Gu, Yiran Ding, Ke Sun, Fang Guo, Panzhong Lu, Zhiyuan Ning, Yixuan Weng and Yue Zhang . . . . .	57
<i>QFinZero: A Unified Financial Toolchain for LLM-Based Trading Agents</i> Haochen Luo, Yifan LI, Ho Tin Ko, An Binh Minh, Junjie Xu, Tang Pok Hin, Wang Chak Wong, Gao Yuan, Zhengzhao Lai, Yuan Zhang and Chen Liu . . . . .	68
<i>dLLM: Simple Diffusion Language Modeling</i> Zhanhui Zhou, Lingjie Chen, Hanghang Tong and Dawn Song . . . . .	78
<i>Fast-MIA: Efficient and Scalable Membership Inference for LLMs</i> Hiromu Takahashi and Shotaro Ishihara . . . . .	89
<i>CodeDistiller: Automatically Generating Code Libraries for Scientific Coding Agents</i> Peter Jansen, Samiah Hassan and Pragnya Narasimha . . . . .	99
<i>IDP Accelerator: Agentic Document Intelligence from Extraction to Compliance Validation</i> Md Mofijul Islam, Md Sirajus Salekin, Joe King, Priyashree Roy, Vamsi Thilak Gudi, Spencer Romo, Akhil Nooney, Bob Strahan, Boyi Xie and Diego A. Socolinsky . . . . .	108
<i>AIInterviewer: A Platform for Designing and Conducting AI-led Qualitative Interviews</i> Tobias Gårdhus, Nikolas Vitsakis, Fie Lejre Frederiksen, Anna Rogers and Hjalmar Bang Carlsen	119
<i>AutoForest: Automatically Generating Forest Plots from Biomedical Studies with End-to-End Evidence Extraction and Synthesis</i> Massimiliano Pronesti, Angelo Miculescu, Mohsin Kapdi, Paul Flanagan, Oisín Redmond, Joao H Bettencourt-Silva, Gurdeep Singh Mannu, Spiros Denaxas, Rui Providencia, Anya Belz and Yufang Hou . . . . .	128
<i>Dash-M5H: An Interactive Dashboard for Multi-Modal, Multi-Model Mental Health Assessment</i> Raymond Alavo, Xinyuan Zhang, Gemza Ademaj, Junhui Cai, Hyeokhyen Kwon, Robert Cotes, Gari D. Clifford and Ahmed Abbasi . . . . .	138

<i>MixtureKit: A General Framework for Composing, Training, and Visualizing Mixture-of-Experts Models</i>	Ahmad Chamma, Omar El Herraoui and Guokan Shang .....	148
<i>Squrve: A Unified and Modular Framework for Complex Real-World Text-to-SQL Tasks</i>	Yihan Wang, Peiyu Liu, Runyu Chen, Jiaying Pu and Wei Xu .....	157
<i>RiskLab: A Controlled Toolkit for Probing Emergent Risks in LLM-Based Multi-Agent Systems</i>	Yu Jiang, Wenjie Wang, Yue Huang, Yanbo Wang, Zhenhong Zhou, Xiuying Chen, Yang Liu, Pin-Yu Chen, Wei Wang and Xiangliang Zhang .....	167
<i>ArcLight: A Lightweight LLM Inference Architecture for Many-Core CPUs</i>	Yuzhuang Xu, Xu Han, Yuxuan Li and Wanxiang Che .....	178
<i>DialogGuard: Multi-Agent Psychosocial Safety Evaluation Interface of Sensitive LLM Responses</i>	Han Luo and Guy Laban .....	187
<i>ClinQueryAgent: A Conversational Agent for Population Health Management</i>	Joseph Spartacus Boyle, Anthony Michael Dranfield, Mike O’Neil, Maria Liakata and Alison Q. Smithard .....	197
<i>Rankify: A Comprehensive Python Toolkit for Retrieval, Re-Ranking, and Retrieval-Augmented Generation</i>	Abdelrahman Abdallah, Bhawna Piryani, Jamshid Mozafari, Andreas Herzinger, Jamie Holdcroft and Adam Jatowt .....	208
<i>ScheMatiQ: From Research Question to Structured Data through Interactive Schema Discovery</i>	Shahar Levy, Eliya Habba, Reshef Mintz, Barak Raveh, Renana Keydar and Gabriel Stanovsky	220
<i>GovScope: A Public Multimodal Search System for 70 Million Pages of Government PDFs</i>	Ying-Hsiang Huang, Claire Gong, Shreya Shaji, Alison R Yan, Leslie Harka, Albert Du, Anjali Shubha Gopal, Samuel J Klein, Shannon Zejiang Shen, Mark E. Phillips, Trevor Owens, Kyle Deeds and Benjamin Charles Germain Lee .....	231
<i>Bridging Kernel Drivers and Virtual Device Models with LLM-Powered Automation</i>	Mingyu Wang, Bin Yu, Wenjian Lu, Zhi Wang, Gao Kefeng, Cheng Wen, Xu Lu and Cong Tian	242
<i>BabelDOC: Better Layout-Preserving PDF Translation via Intermediate Representation</i>	Yang Qi, Xiangyao Ma, Xiao Wang, Hao Wang and Rui Wang .....	253
<i>JointCoder: Exploring Automated ICD Coding on Real-World Chinese EHRs with a Multi-Agent Framework</i>	Kangjun Liu, Zhenyu Li, Jianlei Wang, Hongjiao Guan, Ying Lian, Guoqiang Chen, Tao Xin and Wenpeng Lu .....	263
<i>Hindsight: Structured Agent Memory that Retains, Recalls, and Reflects</i>	Christopher Latimer, Nicolò Boschi, Andrew Neeser, Chris Bartholomew, Gaurav Srivastava, Xuan Wang and Naren Ramakrishnan .....	275
<i>Semantic XPath: Structured Agent Memory Access for Conversational AI</i>	Yifan Simon Liu, Ruifan Wu, Liam Gallagher, Jiazhou Liang, Armin Toroghi and Scott Sanner	286

<i>Graph of Trace: Visualizing Execution Traces of Scientific Agents</i>	
Tianci Gao, Haoxuan Li, Jian He Li, Tianxiang Zhao, Shi Runze, Weiran Wang, Zezhao Wu and Lu Mi .....	297
<i>ScaleBox: Enabling High-Fidelity and Scalable Code Verification for Large Language Models</i>	
Jiasheng Zheng, Xin Zheng, Boxi Cao, Pengbo Wang, Zhengzhao Ma, Qiming Zhu, Jiazhen Jiang, Yaojie Lu, Hongyu Lin, Xianpei Han and Le Sun .....	307
<i>DataArc-SynData-Toolkit: A Unified Closed-Loop Framework for Multi-Path, Multimodal, and Multilingual Data Synthesis</i>	
Zhichao Shi, Cehao Yang, Hao Zhou, Xiaojun Wu, Huajie Li, Xuhui Jiang, Chengjin Xu, Yuanzhuo Wang and Jian Guo .....	318
<i>HuoziIME: An On-Device LLM-Enhanced Input Method for Deep Personalization</i>	
Baocai Shan, Yuzhuang Xu and Wanxiang Che .....	327
<i>WIGVO: Real-Time Bidirectional Speech Translation over Legacy PSTN Calls via Dual-Session Echo Gating</i>	
Hyeong-seob Kim, Sang-Woo Son, Hyun-woo Cho, Hyeonsang Kim and Jinmo Kim .....	336
<i>MASEval: Extending Multi-Agent Evaluation from Models to Systems</i>	
Cornelius Emde, Alexander Rubinstein, Anmol Goel, Ahmed Heakl, Sangdoo Yun, Seong Joon Oh and Martin Gubri .....	345
<i>MASFactory: A Graph-centric Framework for Orchestrating LLM-Based Multi-Agent Systems with Vibe Graphing</i>	
Yang Liu, Jinxuan Cai, Yishen Li, Qi Meng, Zedi Liu, Xin Li, Chen Qian, Chuan Shi and Cheng Yang .....	357
<i>FactSearch: An Interactive Agentic Fact Search System for Verifying Large Language Model Outputs</i>	
Meng Fang and Harry Mackenzie .....	367
<i>Potato 2.0: A Comprehensive Annotation Platform with AI-in-the-Loop Support</i>	
David Jurgens, Michael Chen and Lina Iyer .....	374
<i>mllm-shap: A Shapley Value Explainability Platform for Text-Audio Multimodal Large Language Models</i>	
Jakub Muszyński, Paweł Pozorski and Maria Ganzha .....	387
<i>PaperMentor: A Human-Centered Multi-Agent Writing Tutor for AI Research Papers in Overleaf</i>	
Jiarui Liu, Terry Jingchen Zhang, Ryan Faulkner, Xuanqiang Angelo Huang, Vilém Zouhar, Dominik Glandorf, Isabel Dahlgren, Rishit Dagli, Yuen Chen, Felix Leeb, Van Q. Truong, Punya Syon Pandey, Yves Bicker, Suvajit Majumder, Wenyuan Jiang, Zeju Qiu, Sankalan Pal Chowdhury, Mrinmaya Sachan, Bernhard Schölkopf, Mona T. Diab and Zhijing Jin .....	397
<i>Gardener: An Agentic AI System for Single-Cell RNA Sequence Analysis</i>	
Junhan Liu, Zhenke Liu, Yongcheng Shi, Peilin Yu, Minxing Zhang and Jiapeng Zhang .....	408
<i>TokCollate: A Comprehensive Tool for Tokenizer Evaluation and Visualization across Languages</i>	
Dušan Variš, Abishek Stephen and Jindřich Libovický .....	418
<i>A System for Dynamically Tracking Content Moderation on Reddit</i>	
George Arthur Baker and Bharadwaj Kadiyala .....	428
<i>AI Steerability 360: A Toolkit for Steering Large Language Models</i>	
Erik Miehl, Karthikeyan Natesan Ramamurthy, Praveen Venkateswaran, Ching-Yun Ko, Pierre Dognin, Moninder Singh, Tejaswini Pedapati, Avinash Balakrishnan, Matthew Riemer, Dennis Wei, Inge Vejsbjerg, Elizabeth M. Daly and Kush R. Varshney .....	436

<i>Media-to-Insights: A Multi-Agent AI System for Continuous Media Monitoring, Analysis, and Reporting</i>	
Ashraf Hatim Elneima, Ozan Yilmaz, Hadi Nasrallah, Sanjika Hewavitharana, Mohamed Al-Badrashiny and Hassan Sawaf . . . . .	445
<i>LinkNav: Surfacing Interconnected Information in Scientific Articles</i>	
Sebastian Antony Joseph, Jennifer Healey, Junyi Jessy Li and Ani Nenkova . . . . .	453
<i>FourCorners: Grounded Thai Legal Research over a Temporal Knowledge Graph</i>	
Pawitsapak Akarajaradwong, Thitiwat Nopparatbundit, Treephop Saeteng, Kasidit Phoncharoen, Sarana Nutanong and Chompakorn Chaksangchaichot . . . . .	463
<i>rosaOS: Agentic Operating System for Embodied LLMs</i>	
Yijun Ge, Kushaldeep Mujral, Karthik Nambiar and Jimmy Lin . . . . .	473
<i>CritiSense: Critical Digital Literacy and Resilience Against Misinformation</i>	
Firoj Alam, Fatema Ahmad, Ali Ezzat Shahroor, Mohamed Bayan Kmainasi, Elisa Sartori, Giovanni Da San Martino, Abul Hasnat and Raian Ali . . . . .	481
<i>MED-COPILOT: A Medical Assistant Powered by GraphRAG and Similar Patient Case Retrieval</i>	
Shuheng Chen, Namratha Patil, Haonan Pan, Angel Hsing-Chi Hwang, Yao Du, Ruishan Liu and Jieyu Zhao . . . . .	493
<i>PathBuilder: A Quality-Controlled LLM System for Personalized Learning Pathways</i>	
Jasper Meynard Arana, John Andrew Mañacop, John Allen Manacop, Roy Andrew Garcia, Keith Rick Piniera, Kristine Ann M. Carandang, Ethan Robert Casin, Christian Alis and Christopher Monterola . . . . .	504
<i>AutoChecklist: Composable Pipelines for Checklist Generation and Scoring with LLM-as-a-Judge</i>	
Karen Zhou and Chenhao Tan . . . . .	515
<i>JobMatchAI - An Intelligent Job Matching Platform Using Knowledge Graphs, Semantic Search and Explainable AI</i>	
Mayank Vyas, Abhijit Chakraborty and Vivek Gupta . . . . .	526
<i>Qayyem: A Real-time Platform for Scoring Proficiency of Arabic Essays</i>	
Hoor Tamer Elbahnasawi, Marwan Sayed, Sohaila Eltanbouly, Fatima Zahra Brahamia and Tamer Elsayed . . . . .	536
<i>PhyVer: Physics-Grounded Material Claim Verification with Multi-Fidelity Physical Evidence</i>	
Jianpeng Chen, Wangzhi Zhan, Haohui Wang, Brian Mayer, Dongqi Fu and Dawei Zhou . . . . .	546
<i>pAtChWoRK: Patching the Pieces of Public Procurement Documents</i>	
Lorena Calvo-Bartolomé, Saúl Blanco Fortes, Erick Cedeño and Jerónimo Arenas-García . . . . .	556
<i>UltraEval-Audio: A Unified Framework for Comprehensive Evaluation of Audio Foundation Models</i>	
Qundong Shi, Jie Zhou, Biyuan Lin, Junbo Cui, Guoyang Zeng, Yixuan Zhou, Ziyang Wang, Xin Liu, Zhen Luo, Yudong Wang and Zhiyuan Liu . . . . .	566
<i>Paper2Web: Let's Make Your Paper Alive!</i>	
Yuhang Chen, Tianpeng Lv, Yao Wan, Philip S. Yu and Dongping Chen . . . . .	578
<i>OxyGent: Making Multi-Agent Systems Modular, Observable, and Evolvable via Oxy Abstraction</i>	
Junxing Hu, Tianlong Li, Lei Yu and Ai Han . . . . .	586
<i>EduCoder: An Open-Source Annotation System for Education Transcript Data</i>	
Saad Ashraf, Jim Malamut, Vishal Kumar, Guanzhong Pan, HyunJi Nam, Mei Tan, Lucía Langlois, Liliana Carolina Santos-Deonizio, Helen Spencer Higgins and Dorottya Demszky . . . . .	597

<i>CPTCoder: A Reliable LLM System for Medical Procedure Code Prediction</i> Benlu Wang, Ziyao Shangguan, Kyle Tegtmeier, Zhenyu Zhang, Sophie Chheang and Arman Cohan.....	605
<i>ParseJargon: Personalized Real-time Jargon Support in Online Meetings</i> Yifan Song, Wing Yee Au, Hon Yung Wong, Brian Bailey and Tal August .....	615
<i>The olmOCR Project: Building Fully Open OCR using VLMs</i> Jake Poznanski, Kyle Lo and Luca Soldaini.....	626
<i>AiraXiv: An AI-Driven Open-Access Platform for Human and AI Scientists</i> Junshu Pan, Panzhong Lu, Yixuan Weng, QiYao Sun, Fang Guo, Zijie Yang, Qiji Zhou and Yue Zhang.....	636
<i>TruthSplit: Revealing Conditional Validity in Arguments Through Multi-Worldview Comparative Reasoning</i> Benjamin Stieger, Maximilian Terberger, Thomas Huber and Christina Niklaus .....	648
<i>TONY: an open-source TOOLkit for Nlp in psYchology</i> Federico Ravenda, Sofia Irene Ravenda, Volodymyr Karpenko, Daniele Montagnani, Andrea Raballo and Antonietta Mira.....	660
<i>Formally Specifying the Intended Behavior of the Program: LLM-Driven Neuro-Symbolic Program Specification Synthesis</i> Cheng Wen, Hu Junjie, YiKun Hu, Jie Su, Bin Yu, Dugang Liu, Zhiwu Xu, Weidi Sun, Shengchao Qin and Cong Tian .....	672
<i>AnnoHID: LLM-Assisted Annotation Framework for Low-Resource Medical Texts</i> Annisa Maulida Ningtyas, Guntur Budi Herwanto, Yunita Sari, Rifki Afina Putri, Filip Kovacevic, Alaa El-Ebshihy, Varvara Arzt and Florina Piroi.....	683
<i>RECAP: An End-to-End Platform for Capturing, Replaying, and Analyzing AI-Assisted Programming Interactions</i> Keyu He, Qianou Ma, Valerie Chen, Wayne Chi and Tongshuang Wu .....	692
<i>A Dynamic Self-Evolving Extraction System</i> Moin Aminnaseri, Hannah Kim and Estevam Hruschka .....	702
<i>ThinkBooster: A Unified Framework for Seamless Test-Time Scaling of LLM Reasoning</i> Vladislav Smirnov, Quang-Chieu Nguyen, Sergey Senichev, Minh Ngoc Ta, Ekaterina Fadeeva, Artem Vazhentsev, Daria Galimzianova, Nikolai Rozanov, Viktor Mazanov, Jingwei Ni, Tianyi Wu, Igor Kiselev, Mrinmaya Sachan, Iryna Gurevych, Preslav Nakov, Timothy Baldwin and Artem Shelmanov .....	715
<i>FinReporting: An Agentic Workflow for Localized Reporting of Cross-Jurisdiction Financial Disclosure</i> Fan Zhang, Mingzi Song, Rania Elbadry, Yankai Chen, Shaobo Wang, Yixi Zhou, Xunwen Zheng, Yueru He, Yuyang Dai, Georgi Nenkov Georgiev, Ayesha Gull, Muhammad Usman Safder, Fan Wu, Liyuan Meng, Fengxian Ji, Junning Zhao, Xueqing Peng, Jimin Huang, YU Chen, Xue Liu, Preslav Nakov and Zhuohan Xie .....	728
<i>Expert Calibration Lens for Pruning Mixture of Experts</i> Luis Frentzen Salim, Chia-Chun Wu, Tran Van Nhiem, Lun-Wei Ku and Yung-Hui Li .....	736
<i>Ryze: Evidence-Enriched Data Synthesis from Biomedical Papers</i> Yeqi Huang, Yue Chen, Yanwei Ye, Guanhao Su and Luo Mai .....	743

<i>Agentic CLEAR: Automating Multi-Level Evaluation of LLM Agents</i> Asaf Yehudai, Lilach Eden and Michal Shmueli-Scheuer .....	750
<i>FD-NL2SQL: Feedback-Driven Clinical NL2SQL that Improves with Use</i> Suparno Roy Chowdhury, Tejas Anvekar, Manan Roy Choudhury, Muhammad Ali Khan, Kaneez Zahra Rubab Khakwani, M Bassam Sonbol, Irbaz Bin Riaz and Vivek Gupta .....	764
<i>DIAGRAMS : A Review Framework for Reasoning-Level Attribution in Diagram QA</i> Anirudh Iyengar Kaniyar Narayana Iyengar, Tampu Ravi Kumar, Manan Suri, Raviteja Bom-mireddy, Dinesh Manocha, Puneet Mathur and Vivek Gupta .....	774
<i>SlideGuard: AI-Driven Evaluation of Graduate Student Presentation Materials</i> Nikolay Alekseevich Butakov, Maria Khodorchenko, Mazein Nikita, Daniil Gareev, Yuri Falevskiy, Georgii Konev and Denis Nasonov .....	785
<i>Spectra: A Mechanistic Interpretability Library for Vision-Language Models</i> Clement Neo, Yongsen Zheng, Kwok-Yan Lam and Luke Ong .....	794
<i>Demonstrating ViviDoc: Generating Interactive Documents through Human-Agent Collaboration</i> Yinghao Tang, Yupeng Xie, Yingchaojie Feng, Tingfeng Lan and Wei Chen .....	804
<i>Praat++: Multimedia Annotation System for Speech and Vocalization</i> Weiran Zhang and Kenny Q. Zhu .....	812
<i>AgentFactory: A Self-Evolving Framework Through Executable Subagent Accumulation and Reuse</i> Zhang Zhang, Shuqi Lu, Hongjin Qian, Di He and Zheng Liu .....	819
<i>OpenGlass: A Sensing-Computing Split Architecture for Local MLLM-Driven Real-Time Visual Assis-tance</i> Mengzhang Li and Yuan Yao .....	829
<i>TartanMaroon: Multi-Agent Academic Advising with Iterative Negotiation and Transparent Collabora-tion</i> Peidi Dong, Houda Bouamor, Yunze Xiao and Devi G Kurup .....	840
<i>GAMED.AI: A Hierarchical Multi-Agent Framework for Automated Educational Game Generation</i> Shiven Agarwal, Yash Shah, Ashish Raj Shekhar, Priyanuj Bordoloi and Vivek Gupta .....	851
<i>MOOSE-Copilot: A Web-Based Interactive Assistant for Unified Exploratory and Fine-Grained Scien-tific Hypothesis Discovery</i> Hongran An and Zonglin Yang .....	861