

EMNLP 2025

**The 2025 Conference on Empirical Methods in Natural  
Language Processing**

**Proceedings of the Conference**

November 4-9, 2025

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

## Destination Host



## Diamond



## Platinum



## Gold



JPMorganChase



## Silver



Appen

Bloomberg  
Engineering



## Bronze



keensight



## Diversity and Inclusion: Ally



©2025 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 979-8-89176-332-6

## Message from the General Chair

I am excited to welcome you to this year's edition of the Conference on Empirical Methods in Natural Language Processing! Importantly, it marks the 30th edition of EMNLP. With over 8,000 submissions, more than 3,000 accepted papers, and thousands of attendees, we have come a long way from that first workshop, which had 14 accepted papers. As the field looks ahead, Suzhou is the fitting location for celebrating this milestone: rooted in a long literary tradition, yet modern and forward-looking, and home to a large share of the NLP community.

A conference may seem deceptively straightforward: accept a bunch of papers, make some of them presentations and the rest posters, and done! Far from it, though. As the numbers already suggest, the logistics behind EMNLP have become highly complex. Apart from ACL's professional conference organization team around the indefatigable Jenn Rachford, Lina Staggs, Megs Haddad, and Damira Mrsic from Underline, it took an army of dedicated volunteers coordinating messages, room assignments, acceptance decisions, and dealing with last-minute requests, complaints, and the administrative, financial, and bureaucratic minutiae of getting thousands of people from across the world to get together in Suzhou.

My heartfelt thanks go to the entire organization committee and their tireless efforts. Without them and their relentless dedication, this conference would not have been possible. If you see one of them, please do thank them.

- **Program Chairs:** Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, Violet Peng
- **ARR Guest Program Chair:** Sarvnaz Karimi, Margot Mieskes, Michael White
- **Local Chairs:** Wei Ji, Jiaheng Liu
- **Industry Track Chairs:** Saloni Potdar, Lina Rojas-Barahona, Sebastien Montella
- **Workshop Chairs:** Sunipa Dev, Maja Popović, Eleftherios Avramidis
- **Tutorial Chairs:** Valentina Pyatkin, Andreas Vlachos
- **Virtual Infrastructure Chair:** Lianhui Qin
- **Sponsorship Chairs:** Joel Tetreault, Mengwen Liu
- **Ethics Chairs:** Shashank Srivastava, Anil Ramakrishna
- **Internal Communications Chairs:** Juntao Li, Steve Wilson
- **Demonstration Chairs:** Ivan Habernal, Jörg Tiedemann, Peter Schulam
- **Publication Chairs:** Shivashankar Subramanian, Hyunsoo Cho, Mark Hopkins
- **Best Publication Chairs:** Mirella Lapata, Owen Rambow
- **Handbook Chairs:** Lucia Siciliani
- **Publicity Chairs:** Yash Kumar Lal, Aaron Mueller, Dallas Card
- **Student Volunteer Chairs:** Nadjma Ousidhoum, Anne Lauscher
- **Diversity/Inclusion Chairs:** Chenxi Whitehouse, Margot Mieskes, Roy Ka-Wei Lee
- **Website Chairs:** Marcelo Viridiano, Haiyue Song
- **Virtual Infrastructure Chairs:** Lianhui Qin

And of course, to the ACs, SACs, and countless reviewers who made the scientific program possible. Thank you also to MCI, our local conference team, and the SIGDAT board: Raquel Fernandez, Isabelle Augenstein, Kai-Wei Chang, Alice Oh, and Juan Pino, for all their support and trust. Special thanks to David Yarowsky, one of the founders of EMNLP, and Yuji Matsumoto for their perspective on the conference's history and their continued support.

I am grateful to the PCs who tackled the new challenges of growing submission numbers, AI-generated text, and addressing missing reviewers with a range of successful initiatives. It's a sign of a resilient and adaptive field moving forward.

I hope this conference will leave you feeling inspired, motivated, and energized to return to your work after a wonderful time in Suzhou. Happy birthday, EMNLP: here's to many more editions!

Dirk Hovy  
*EMNLP 2025 General Chair*

## Message from the Program Chairs

Welcome to the 2025 Conference on Empirical Methods in Natural Language Processing, which is the 30th EMNLP. The field has a lot to be proud of, looking back, and not just about recent advances. A lot has changed across these 30 years. The field of NLP has arguably ushered in the era of Generative AI, transforming itself in the process. Even before LLMs however, EMNLP has witnessed a massive growth, going from a small gathering with 14 accepted papers to a major event with thousands of accepted papers. EMNLP 2025 will be held in a hybrid format, offering attendees the option to join us in person in beautiful Suzhou, China, or to participate remotely from anywhere in the world. We are grateful to be able to host this year's conference in such a historic site, noted as one of the most beautiful cities in China, and famous for intricate, lovely embroidery. Like with every past conference, the work over the past year organizing EMNLP 2025 has been a collaborative effort, made possible by the dedication and hard work of thousands of people. We would like to thank the support and contributions of the following people:

- The General Chair, Dirk Hovy
- The ARR Editors-in-Chief of the May 2025 cycle (Sarvnaz Karmi, Margot Mieskes, & Michael White), Technical Staff (Sudipta Kar, Freda Shi, Jonathan Kummerfeld, Yoshimoto Matsubara, Dhruv Naik, and Holy Lovenia), and the extended team
- The OpenReview team, especially Harold Rubio and Melisa Bok
- The 168 Senior Area Chairs
- The 1,989 Area Chairs and the 13,048 reviewers
- The awards committee chairs, Mirella Lapata and Owen Rambow, and the awards committee members
- The ethics chairs, Shashank Srivastava and Anil Ramakrishna
- The industry track chairs, Saloni Potdar, Lina Rojas-Barahona, and Sebastian Montella
- The demonstration chairs, Ivan Habernal, Joerg Tiedemann, and Peter Schulam
- The internal communications chairs, Juntao Li and Steve Wilson
- The website chairs, Marcelo Viridiano and Haiyue Song
- The publication chairs, Shivashankar Subramanian, Hyunsoo Cho, and Mark Hopkins
- The handbook chair, Lucia Siciliani
- The local organization chairs, Wei Ji and Jiaheng Liu, and their team
- The publicity chairs, Yash Kumar Lal, Aaron Mueller, and Dalas Card
- The student volunteer chairs, Nedjma Ousidhoum and Anne Lauscher
- The diversity/inclusion chairs, Chenxi Whitehouse, Margot Mieskes, and Roy Ka-Wei Lee
- The virtual infrastructure chair, Lianhui Qin
- The workshop chairs, Sunipa Dev, Maja Popovic, and Eleftherios Avranidis
- The tutorial chairs, Valentina Pyatkin and Andreas Vlachos

- The sponsorship chairs, Joel Tetreault and Mengwen Liu
- The ACL Anthology Director Matt Post and his team
- The TACL editors-in-chief (Asli Celikyilmaz, Roi Reichart, and Dilek Hakkani Tur) and CL Editor in-Chief Wei Lu
- The NAACL 2024 Program Chairs (Kevin Duh, Helena Gomez, and Steven Bethard) and the ACL 2024 Program Chairs (Lun-Wei Ku, Andre F. T. Martins, Vivek Srikumar)
- Damira Mrsic and Underline Team
- Jennifer Rachford and conference support staff
- All the authors of papers submitted for review

## Review Process

All submissions to EMNLP 2025 went through a two-stage review process. First, papers were submitted to the ACL Rolling Review (ARR), where they were reviewed by reviewers and received meta-reviews from area chairs. Authors then had the option to commit their reviewed papers to EMNLP via a separate EMNLP 2025 commitment site. At this stage, senior area chairs provided recommendations, and final acceptance decisions were made by the program chairs. This process is consistent with previous conferences, including NAACL 2025 and ACL 2025, and others. We worked closely with the ARR team, particularly the Editors-in-Chief for the May 2025 cycle, and served as guest Editors-in-Chief for this round. We also recruited additional reviewers and area chairs for ARR, bringing the total to 13,048 reviewers and 1,989 area chairs in the May 2025 ARR cycle, which handled the majority of EMNLP 2025 submissions. EMNLP also recruited 168 senior area chairs to oversee the review and meta-review process.

## New Review Incentives Policy

This year, in collaboration with ARR, we implemented a new policy to help us meet the ever growing reviewing demands of the \*CL conferences. In order to meet reviewer demand, it is necessary to require reviewing from a broader range of authors, which in turn creates challenges in maintaining a high quality standard as well as ensuring that reviews are submitted in a timely fashion. You can read the full policy at <https://aclrollingreview.org/incentives2025> but the most important element was that reviewers who were deemed “highly irresponsible” would not be allowed to commit their work to EMNLP or (re-)submit their work to the July ARR cycle. This year, we were cautious in implementing the policy, only desk-rejecting (and disallowing submission to the July ARR cycle) papers where at least one author met all the following criteria:

- They were assigned papers to review at the beginning of the review cycle (i.e. not emergency reviews).
- They were a reviewer with at least 3 assigned papers and failed to submit 100% of their reviews despite multiple reminders.
- They did not submit an Emergency declaration or a Delay notification.

Out of the 13,048 reviewers recruited this year, only 69 were deemed “highly irresponsible” under our standards, and enforcement was applied solely in those cases. In future cycles, this effort will extend more broadly to include Area Chairs, and reviewers with a reduced workload, or those who did some portion, but not enough, of what they were assigned. We will also be implementing a warning system



where co-authors would be informed about any of their co-authors that meet the “highly irresponsible” reviewer criteria. This was piloted in the July ARR cycle. Overall, the ARR process ran somewhat less smoothly than ideal, with more than 50% of submitted papers missing at least one review by the deadline. We worked as a team to ensure that all submitted papers received at least three reviews and a meta-review, which required rounds of reminder emails and invitation of emergency reviewers. For the EMNLP commitment phase, senior area chairs made recommendations for 4512 committed papers based on the reviews, meta-reviews, and the papers themselves, with final acceptance decisions made by the program chairs.

## Acceptance Rate

In total, there are 1811 papers accepted to the Main Conference and 1417 papers accepted to Findings. The acceptance rate for Main Conference papers is 22.16% and one for Findings papers is further 17.34%. The acceptance rate calculation largely follows precedent set by previous conferences that also go through ACL Rolling Review (ARR), e. g. NAACL 2025, ACL 2025. The calculation takes into account the multi-stage process of ARR, where a paper may get revised in ARR and then later committed to the conference. The denominator includes all those committed to EMNLP from Dec 2024 ARR through May 2025 ARR.

|                          |      |
|--------------------------|------|
| Papers from ARR May 2025 | 7916 |
| Papers from ARR Feb 2025 | 220  |
| Papers from ARR Dec 2024 | 25   |
| Papers from ARR Oct 2024 | 11   |
| Papers from ARR Aug 2024 | 1    |
| Papers from ARR Jun 2024 | 1    |

Among these submissions, 778 were withdrawn before receiving three reviews, and 670 were desk-rejected. These are still counted in the denominator. The acceptance rate for Main Conference papers is therefore:  $1811/8174 = 22.16\%$ . The final Main Conference proceedings will include 1668 long papers and 143 short papers. Findings papers are those which are not accepted at the Main Conference, but nevertheless have been judged worthy of publication as “solid work with sufficient substance, quality and novelty”. The next  $1417/8174 = 17.34\%$  of papers were accepted to EMNLP Findings. The final Findings proceedings will include 1246 long papers and 149 short papers.

## Special Theme: Advancing our Reach – Interdisciplinary Recontextualization of NLP

The core interests of the ACL community are rooted in human-language technologies but also have a broad reach into other fields. A couple of recent examples are the burgeoning areas of code models and vision models. Earlier cases are exemplified through SIGs connected with the fields of education, medicine, and humanities. Movements such as NLP for Social Good and Computational Social Science show a desire for broad impact, which requires expertise beyond the borders of our own community to achieve. This year’s theme of Advancing our Reach: Interdisciplinary Recontextualization of NLP aims to highlight this need for broader connections with other fields to understand and intensify NLP’s impact. The goal is to increase our awareness of how advances in NLP can impact other fields, and design better strategies to measure that impact both within and across disciplines. Over the past two decades, the field has advanced at an exponential rate. The term language models is now a household word, the industry is booming, and the publication rate is dizzying, but what does that mean about fundamental scientific impact and broader impact on real societal problems? How can we measure that in a rigorous way? Scores on benchmarks are increasing; however, to what extent do our benchmarks reflect the true

impact of our technology advances? If we make a distinction between impact within our own field versus impact from our field into other fields, would we see the same magnitude of growth? The conventional measures of success don't facilitate making critical distinctions, like incremental improvement versus transformative change, or within-field uptake versus broad impact across fields. So this year, we invited engagement with the theme first through a theme-specific submission track for papers addressing the fundamental technology advances and papers addressing the evaluation methodology issues. Note that we did not consider position papers. However, we also invited Fireside Chat session proposals designed to bring together NLP researchers with leaders from other fields for agenda setting and new collaboration formation. Finally, we invited multi-disciplinary panel proposals that provide opportunities to engage the broader community in reflection related to the theme. We received 100 submissions to the theme track during the commitment phase. Among these, 41 papers were accepted to the main conference and a further 32 to Findings of EMNLP.

## Best Paper Selection

In keeping with the practices observed in EMNLP 2024, we included the following best paper awards categories:

- **Best Papers** (featuring  $< 0.25\%$  of accepted papers) and **Outstanding Papers** (featuring  $< 2.5\%$  of accepted papers) present fascinating, controversial, surprising, impressive, and/or potentially field-changing ideas.
- **Senior Area Chairs' awards** are similar to Outstanding papers, but specific to this research track.
- **Best Theme Papers** (= Senior Area Chair's awards for the special theme track) make significant new contributions to efficiency in model algorithms, training, and inference.
- **Social Impact Papers** have the potential for a significant positive societal impact.
- **Resource Papers** announce, describe, and share a fascinating, valuable, or potentially field-changing new resource.

Based on nominations from SACs and ACs, 75 candidates were shortlisted for consideration for the above awards. The final selection was made by the Best Paper Award Committee, and the winners will be announced and will present their work during the closing ceremony..

## Program Composition & Presentation Modes

Based on feedback recommendations from the General Chair and Jennifer Rachford's team, we decided to hold the virtual poster sessions immediately before and after the in-person program on the main conference days. This approach allows virtual attendees to participate concurrently with the physical event, avoiding the need for organizers and attendees to engage with the conference twice. This year, 325 main conference papers were selected for oral presentations by the program chairs, with the goal of creating a well-rounded program featuring a diverse set of topics instead of selecting papers based on their review scores. Numerical scores were only part of the consideration of assignment to oral versus poster presentations. We placed a high value on representation across tracks and topical coherence in the sessions. In addition to the main conference papers, the EMNLP program also includes 2 papers accepted by Computational Linguistics and 15 papers accepted by Transactions of the ACL (TACL). Among these, 17 journal papers will be presented in-person as oral presentations, thematically distributed across appropriate sessions, with formats chosen based on author preferences. Additionally, all Findings papers have been assigned poster presentations in designated sessions. Rounding out the program are dedicated sessions for the demonstrations track and the student research workshop.

## Keynotes and Panel

This year's program features an impressive lineup of three keynote speakers:

- Heng Ji, University of Illinois at Urbana-Champaign, USA
- Jana Diesner, Technical University of Munich, Germany
- Hannaneh Hajishirzi, University of Washington, USA

Alongside these keynotes, we are looking forward to hosting a panel discussion related to the conference theme entitled “*Advancing our Reach: Do's, Don'ts and Lessons learned in Interdisciplinary Research*”, which will include the keynote speakers and other invited guests.

We hope that you will enjoy this year's program and hybrid conference!

Christos Christodoulopoulos, Information Commissioner's Office (ex-Amazon)

Tanmoy Chakraborty, Indian Institute of Technology Delhi (IIT Delhi)

Carolyn Rose, Carnegie Mellon University (CMU)

Violet Peng, University of California, Los Angeles (UCLA)

*EMNLP 2025 Program Co-Chairs*

## Message from the Local Chairs

Dear EMNLP 2025 Participants,

It is our great honor to welcome you to EMNLP 2025 in Suzhou, China — a city where classical elegance meets modern innovation. Often called the "Venice of the East", Suzhou has for centuries inspired poets, scholars, and artists with its winding canals, stone bridges, and UNESCO World Heritage classical gardens. Today, it also stands as a vibrant hub of research, technology, and international exchange, making it a fitting setting for one of the premier conferences in natural language processing.

Suzhou is a city of languages in every sense. The mellifluous tones of the Wu dialect, once praised as the "language of songs", mingle with Mandarin and the many tongues of visitors and residents from around the world. This rich linguistic and cultural landscape reflects the very spirit of EMNLP: embracing diversity, fostering dialogue, and advancing our shared understanding of human and machine language.

Beyond its cultural heritage, Suzhou has rapidly emerged as a center of science and innovation. Anchored by world-class universities, thriving industries, and close connections to Shanghai, it embodies the dynamic transformation that continues to shape both China and the global community. In this blend of tradition and modernity, we see a metaphor for our field: deeply rooted in language's long history, yet continually branching into new technologies and applications.

During your stay, we invite you to explore the beauty and vitality of Suzhou. Stroll through the Humble Administrator's Garden or the Lingering Garden, where architecture and nature merge in timeless harmony. Wander the ancient water towns of Zhouzhuang and Tongli, where history flows quietly along canals. Discover the contemporary skyline of Jinji Lake, or visit the Suzhou Museum, masterfully designed by I. M. Pei. And of course, delight in Suzhou's culinary traditions, from the artistry of Suzhou-style noodles to the elegance of seasonal delicacies.

We extend our deepest gratitude to the organizing team, and the many volunteers whose dedication has made this conference possible.

We hope that EMNLP 2025 will be both intellectually stimulating and personally memorable, offering you not only a rich technical program but also the chance to experience the unique charm of Suzhou — a city where water, garden, and language flow together in harmony.

Welcome to Suzhou!

Wei Ji, Nanjing University  
Jiaheng Liu, Nanjing University  
*EMNLP 2025 Local Chairs*

# Organizing Committee

## General Chair

Dirk Hovy, Bocconi University

## Program Chairs

Christos Christodoulopoulos, Information Commissioner's Office

Tanmoy Chakraborty, Indian Institute of Technology Delhi

Carolyn Rose, Carnegie Mellon University

Violet Peng, University of California, Los Angeles

## Guest Program Chairs

Sarvnaz Karimi, CSIRO

Margot Mieskes, University of Applied Sciences, Darmstadt

Michael White, Ohio State University

## Local Organizing Chairs

Wei Ji, Nanjing University, Suzhou Campus

Jiaheng Liu, Nanjing University, Suzhou Campus

## Industry Track Chairs

Saloni Potdar, Apple Inc.

Lina Rojas-Barahona, Orange Labs

Sebastien Montella, Huawei

## Workshop Chairs

Sunipa Dev, Google Research

Maja Popović, Dublin City University and IU University, Berlin

Eleftherios Avramidis, DFKI, Berlin

## Tutorial Chairs

Valentina Pyatkin, Allen Institute for AI; University of Washington

Andreas Vlachos, University of Cambridge

## Sponsorship Chairs

Joel Tetreault, Dataminr, Inc.

Mengwen Liu, Amazon

## Ethics Chairs

Shashank Srivastava, University of North Carolina at Chapel Hill

Anil Ramakrishna, Meta

### **Internal Communications Chairs**

Juntao Li, Soochow University  
Steve Wilson, University of Michigan-Flint

### **Demonstration Chairs**

Ivan Habernal, RC Trust; Ruhr University Bochum  
Jörg Tiedemann, University of Helsinki  
Peter Schulam, Amazon AGI

### **Publication Chairs**

Shivashankar Subramanian, Amazon  
Hyunsoo Cho, Ewha Womans University  
Mark Hopkins, Williams College

### **Best Publication Chairs**

Mirella Lapata, University of Edinburgh  
Owen Rambow, Stony Brook University

### **Handbook Chair**

Lucia Siciliani, University of Bari Aldo Moro

### **Publicity Chairs**

Yash Kumar Lal, Stony Brook University  
Aaron Mueller, Boston University  
Dallas Card, University of Michigan

### **Student Volunteer Chairs**

Nedjma Ousidhoum, Cardiff University  
Anne Lauscher, University of Hamburg

### **Diversity and Inclusion Chairs**

Chenxi Whitehouse, Meta AI  
Margot Mieskes, University of Applied Sciences Darmstadt  
Roy Ka-Wei Lee, Singapore University of Technology and Design

### **Website Chairs**

Marcelo Viridiano, Case Western Reserve University  
Haiyue Songs, National Institute of Information and Communications Technology

# Program Committee

## AI/LLM Agents

Ahmad Beirami  
Arman Cohan  
Libo Qin  
Mengjie Zhao  
Pan Lu  
Sarath Chandar  
Shumin Deng  
Tuo Zhao  
Xuan Wang  
Xuebo Liu  
Yu Su  
Zilong Zheng

## Code Models

Christo Kirov  
Frank F. Xu  
Graham Neubig  
Lingming Zhang  
Md Rizwan Parvez  
Terry Yue Zhuo  
Ziyang Luo

## Computational Social Science, Cultural Analytics, and NLP for Social Good

Ashiqur R. KhudaBukhsh  
David Bamman  
Debora Nozza  
Diyi Yang  
Pawan Goyal  
Preslav Nakov  
Roy Ka-Wei Lee  
Tanu Mitra  
Usman Naseem

## Dialogue and Interactive Systems

Lara J. Martin  
Minlie Huang  
Prithviraj Ammanabrolu

## Discourse, Pragmatics, and Reasoning

Guanyi Chen  
Jad Kabbara

Juntao Yu  
Maciej Ogrodniczuk  
Michael Strube  
Vincent Ng

### **Ethics, Bias, and Fairness**

Hadi Amiri  
Mengnan Du  
Monojit Choudhury  
Natalie Schluter  
Sunipa Dev

### **Generalizability and Transfer**

Adina Williams  
Roi Reichart  
Shay B. Cohen  
Swabha Swayamdipta  
Yue Zhang  
Zoey Liu

### **Hierarchical Structure Prediction, Syntax, and Parsing**

Djamé Seddah  
Kewei Tu  
Lingpeng Kong

### **Human-AI Interaction/Cooperation**

Hao Wang  
Lemei Zhang  
Luciana Benotti  
Lucie Flek  
Philippe Laban  
Tongshuang Wu

### **Information Extraction and Retrieval**

Farah Benamara  
Francis Ferraro  
Krisztian Balog  
Qifan Wang  
Sourish Dasgupta  
Sumit Bhatia  
Wenhu Chen

### **Interpretability, Model Editing, Transparency, and Explainability**

Chenhao Tan  
Gjergji Kasneci



Nadir Durrani  
Peter Hase

### **Linguistic theories, Cognitive Modeling and Psycholinguistics**

Aline Villavicencio  
Cassandra L. Jacobs  
Rui Wang  
Tal Linzen

### **LLM Efficiency**

Akshay Nambi  
Devamanyu Hazarika  
Eduardo Blanco  
Soujanya Poria

### **Low-resource Methods for NLP**

Mehdi Rezagholizadeh  
Miryam de Lhoneux  
Shafiq Joty  
Wasi Uddin Ahmad  
Yue Dong

### **Machine Translation**

Asif Ekbal  
Derek F. Wong  
Sandipan Dandapat

### **Mathematical, Symbolic, and Logical Reasoning in NLP**

Junjie Hu  
Wenpeng Yin  
Zhiting Hu

### **Multilinguality and Language Diversity**

Antonios Anastasopoulos  
Ivan Vulić  
Sebastian Ruder

### **Multimodality and Language Grounding to Vision, Robotics and Beyond**

Huaxiu Yao  
Manling Li  
Xin Eric Wang  
Yonatan Bisk  
Yujun Cai

## **Natural Language Generation**

Kartik Goyal  
Lianhui Qin  
Mohit Iyyer  
Yangfeng Ji

## **Neurosymbolic approaches to NLP**

Fabio Massimo Zanzotto  
Leonardo Ranaldi

## **NLP Applications**

Aakanksha Naik  
Avirup Sil  
Gabriele Pergola  
Gholamreza Haffari  
Gokhan Tur  
Kevin Small  
Makoto Miwa  
Nikolaos Aletras  
Parisa Kordjamshidi  
Roman Klinger  
Sophia Ananiadou  
Sudha Rao  
Tristan Naumann  
Wei Lu  
Yu Zhang

## **Phonology, Morphology and Word Segmentation**

Eleanor Chodroff  
Micha Elsner  
Richard Sproat  
Sarah Moeller

## **Question Answering**

Alessandro Moschitti  
AGI Andrew Yates  
Huan Sun  
Mrinmaya Sachan  
Qingqing Cao  
Yuval Pinter

## **Resources and Evaluation**

Adaku Uchendu  
Ali Emami  
David Ifeoluwa Adelani  
Marzena Karpinska

Massimo Poesio  
Mukul Singh  
Peng Liu  
Salam Khalifa  
Shuly Wintner  
Xi Ye  
Yinya Huang

### **Retrieval-Augmented Language Models**

Jiaming Shen  
Sewon Min  
Yiwei Wang

### **Safety and Alignment in LLMs**

Greg Durrett  
Kai-Wei Chang  
Leon Derczynski  
Ninareh Mehrabi  
Zhuosheng Zhang

### **Semantics: Lexical, Sentence-level Semantics, Textual Inference and Other areas**

Dani Yogatama  
Ellie Pavlick  
Joo-Kyung Kim  
Kuan-Hao Huang  
Teruko Mitamura  
Tristan Miller  
Yi Zhou

### **Sentiment Analysis, Stylistic Analysis, and Argument Mining**

Henning Wachsmuth  
Saif M. Mohammad  
Veronique Hoste

### **Special Theme: Interdisciplinary Recontextualization of NLP**

Armineh Nourbakhsh  
David Jurgens  
Denis Newman-Griffis  
Jana Diesner  
Victoria Yaneva

### **Speech Processing and Spoken Language Understanding**

Frank Rudzicz  
Julia Hirschberg  
Preethi Jyothi

Rolando Coto-Solano  
Sakriani Sakti

## Summarization

Lu Wang  
Sachin Kumar  
Snigdha Chaturvedi

## Area Chairs

Mohamed Abdalla, Ibrahim Abdelaziz, Tarek Abdelzaher, Idris Abdulmumin, Omri Abend, Gavin Abercrombie, Heike Adel, Rodrigo Agerri, Ameeta Agrawal, Sweta Agrawal, Ibrahim Said Ahmad, Alham Fikri Aji, Yamen Ajjour, Alan Akbik, Md Shad Akhtar, Jesujoba Oluwadara Alabi, Firoj Alam, Mohammad Aliannejadi, Emily Allaway, Liesbeth Allein, Milad Alshomary, Fernando Alva-Manchego, Jacob Andreas, Alan Ansell, Xiang Ao, Yuki Arase, Flor Miriam Plaza del Arco, Catherine Arnett, Masaki Asada, Ehsaneddin Asgari, Elliott Ash, Giuseppe Attanasio, Giuseppe Attardi, Katherine Atwell, Dennis Aumiller

Fan Bai, Jiaxin Bai, Long Bai, Xuefeng Bai, JinYeong Bak, Ananth Balashankar, Niranjana Balasubramanian, Timothy Baldwin, David Bamman, Sivaji Bandyopadhyay, Xiaoyi Bao, Yu Bao, Roy Bar-Haim, Chitta Baral, Fazl Barez, Gianni Barlacchi, Jeremy Barnes, Loïc Barrault, Valentin Barriere, Alberto Barrón-Cedeño, Max Bartolo, Farima Fatahi Bayat, Tilman Beck, Amrit Bedi, Yonatan Belinkov, Yassine Benajiba, Adrian Benton, Leon Bergen, Yevgeni Berzak, Steven Bethard, Aditya Bhargava, Victoria W. Bi, Federico Bianchi, Chris Biemann, Jeff Bilmes, Yonatan Bitton, Johannes Bjerva, Terra Blevins, Su Lin Blodgett, Jelke Bloem, Ben Bogin, Danushka Bollegala, Florian Boudin, Zied Bouraoui, Kevin Bowden, Jordan Lee Boyd-Graber, Stephanie Brandl, Pavel Braslavski, Chloe Braud, Davide Buscaldi, Jan Buys, Necva Bölücü

Sky CH-wang, Pere-Lluís Huguet Cabot, Elena Cabrio, Deng Cai, Xiangrui Cai, Nitay Calderon, Iacer Calixto, Jon Ander Campos, Jie Cao, Meng Cao, Pengfei Cao, Qingqing Cao, Shuyang Cao, Yixin Cao, Yong Cao, Ziqiang Cao, Doina Caragea, Dallas Card, Samuel Carton, James Caverlee, Asli Celikyilmaz, Dumitru Clementin Cercel, Tanise Ceron, Soumen Chakrabarti, Chun Kit Chan, Baobao Chang, Kai-Wei Chang, Yi Chang, Shubham Chatterjee, Chung-Chi Chen, Guanhua Chen, Hsin-Hsi Chen, Hui Chen, Jiaao Chen, Jiangjie Chen, Jiaoyan Chen, Jifan Chen, Jinghui Chen, Junfan Chen, Kehai Chen, Lu Chen, Max Chen, Mingda Chen, Muhao Chen, Pin-Yu Chen, Qian Chen, Qianglong Chen, Vivian Chen, Wei Chen, Wenliang Chen, Xilun Chen, Xinchu Chen, Xinyun Chen, Xiuying Chen, Yangyi Chen, Yidong Chen, Yiming Chen, Yulong Chen, Yunmo Chen, Zhaorun Chen, Zhiyu Chen, Zhuang Chen, Fei Cheng, Gong Cheng, Liying Cheng, Lu Cheng, Pu-Jen Cheng, Wei Cheng, Xueqi Cheng, Yu Cheng, Anshuman Chhabra, Hyunsoo Cho, Sangwoo Cho, Edward Choi, Eunsol Choi, YunSeok Choi, Jaegul Choo, Prafulla Kumar Choubey, Jishnu Ray Chowdhury, Chenhui Chu, Simon Clematide, Shay B Cohen, William W. Cohen, Simone Conia, John M Conroy, Caio Corro, Ryan Cotterell, Maxwell Crouse, Heriberto Cuayáhuatl, Leyang Cui, Ruixiang Cui, Yiming Cui, Aron Culotta

Raj Dabre, Xiang Dai, Bhavana Dalvi, Fahim Dalvi, Sandipan Dandapat, Cristian Danescu-Niculescu-Mizil, Amitava Das, Pradeep Dasigi, Pieter Delobelle, Vera Demberg, Daryna Dementieva, Chenlong Deng, Shumin Deng, Yang Deng, Yuntian Deng, Leon Derczynski, Daniel Deutsch, Chris Develder, Barry Devereux, Kuntal Dey, Shizhe Diao, Bosheng Ding, Haibo Ding, Kaize Ding, Liang Ding, Liviu P Dinu, Ajay Divakaran, Lucia Donatelli, Hanze Dong, Li Dong, Ruihai Dong, Zhicheng Dou, Doug Downey, Markus Dreyer, Rotem Dror, Li Du, Mengnan Du, Xinya Du, Brian DuSell, Xiangyu Duan, Nadir Durrani, Ondrej Dusek, Nouha Dziri

Lukas Edman, Liat Ein-Dor, Asif Ekbal, Aparna Elangovan, Maha Elbayad, Ahmed Elgohary, Michael Elhadad, Beyza Ermis, Ori Ernst, Carlos Escolano, Ramy Eskander, Cristina España-Bonet, Luis Espinosa-Anke, Ralph Ewerth

Alexander Richard Fabbri, Hossein Rajaby Faghihi, Agnieszka Falenska, Kai Fan, Wenqi Fan, Yaxin Fan, Biaoyan Fang, Meng Fang, Benamara Farah, Hao Fei, Naomi Feldman, Duanyu Feng, Fuli Feng, Jiazhan Feng, Shangbin Feng, Shaoxiong Feng, Shi Feng, Xiachong Feng, Xiaocheng Feng, Yi Feng, Yue Feng, Eric Le Ferrand, Javier Ferrando, Simone Filice, Mark Fishel, Jeffrey Flanigan, Eve Fleisig, Radu Florian, Antske Fokkens, Mary Ellen Foster, Andre Freitas, Lea Frermann, Daniel Fried, Jinlan Fu, Yanjie Fu, Zihao Fu, Pascale Fung, Yi Fung

Saadia Gabriel, Marco Gaido, Matthias Gallé, Leilei Gan, Zhe Gan, Niloy Ganguly, Pengzhi Gao, Yang Gao, Yifan Gao, Ze-Feng Gao, Judith Gaspers, Jon Gauthier, Jiahui Geng, Luke Gessler, Mor Geva, Abbas Ghaddar, Marjan Ghazvininejad, Deepanway Ghosal, Tirthankar Ghosal, Debanjan Ghosh, Saptarshi Ghosh, Salvatore Giorgi, Mario Giulianelli, Goran Glavaš, Seraphina Goldfarb-Tarrant, Omer Goldman, Dan Goldwasser, Sharon Goldwater, Rob van der Goot, Sruthi Gorantla, Kyle Gorman, Matthew R. Gormley, David Grangier, Loïc Grobol, Jiuxiang Gu, Yu Gu, Lin Gui, Tao Gui, Gael Guibon, Jiafeng Guo, Yue Guo, Zhijiang Guo, Raghav Gupta, Vivek Gupta, Izzeddin Gur, Carlos Gómez-Rodríguez

Nizar Habash, Barry Haddow, Michael Hahn, Dilek Hakkani-Tur, Chi Han, Rujun Han, Xianpei Han, Viktor Hangya, Yiding Hao, Daniel Hardt, David Harwath, Mohammed Hasanuzzaman, Tatsunori Hashimoto, Annette Hautli-Janisz, Shirley Anugrah Hayati, Ben He, Hangfeng He, He He, Jie He, Junxian He, Keqing He, Liang He, Shizhu He, Tianxing He, Wei He, Yulan He, Michael A. Hedderich, Daniel Hershcovich, Toshio Hirasawa, Julia Hockenmaier, Valentin Hofmann, Christopher Homan, Lingzi Hong, Enamul Hoque, Veronique Hoste, Andreas Hotho, Lei Hou, Yufang Hou, Alexander Hoyle, Cho-Jui Hsieh, I-Hung Hsu, Dou Hu, Junjie Hu, Linmei Hu, Mengting Hu, Wei Hu, Xuming Hu, Ziniu Hu, Wenye Hua, Xinyu Hua, Chen Huang, Kung-Hsiang Huang, Ruihong Huang, Shaohan Huang, Tenghao Huang, Ting-Hao Huang, Xinting Huang, Xuanjing Huang, Yi Huang, Yichong Huang, Yinya Huang, Zhenya Huang, Seung-Won Hwang, Wonseok Hwang

Timour Igamberdiev, Oana Ignat, Sathish Reddy Indurthi, Naoya Inoue, Hayate Iso, Masaru Iso-numa

Cassandra L. Jacobs, Kokil Jaidka, Shoaib Jameel, Abhik Jana, Hyeju Jang, Adam Wladyslaw Jatowt, Ganesh Jawahar, Akshita Jha, Harsh Jhamtani, Shaoxiong Ji, Tao Ji, Yangfeng Ji, Robin Jia, Ruoxi Jia, Zixia Jia, Chao Jiang, Feng Jiang, Huiqiang Jiang, Tianyu Jiang, Wenbin Jiang, Yichen Jiang, Yong Jiang, Zhuoxuan Jiang, Yu Jianxing, Allan Jie, Bernal Jimenez, Qiao Jin, Qin Jin, Wei Jin, Zhijing Jin, Xu Jinan, Yohan Jo, Richard Johansson, Aditya Joshi, Wei Ju, Jaap Jumelet, Preethi Jyothi

Jad Kabbara, Hour Kaing, Anup K. Kalia, Ehsan Kamalloo, Hidetaka Kamigaito, Masahiro Kaneko, Dongyeop Kang, Diptesh Kanojia, Börje Karlsson, Pei Ke, Zixuan Ke, Frank Keller, Yova Kementchedjhi, Daniel Khashabi, Khalid Al Khatib, Omar Khattab, Sopan Khosla, Ashiqur R. KhudaBukhsh, Hyunjae Kim, Hyunwoo Kim, Joo-Kyung Kim, Najoung Kim, Taeuk Kim, Yoon Kim, Yunsu Kim, Hirokazu Kiyomaru, Dietrich Klakow, Dan Klein, Mateusz Klimaszewski, Ekaterina Kochmar, Christine De Kock, Tom Kocmi, Philipp Koehn, Konstantinos Kogkalidis, Keshav Kolluru, Rik Koncel-Kedziorski, Lingpeng Kong, Ioannis Konstas, Michalis Korakakis, Fajri Koto, Venelin Kovatchev, Julia Kreutzer, Amrith Krishna, Kalpesh Krishna, Kundan Krishna, Reno

Kriz, Udo Kruschwitz, Shankar Kumar, Shivani Kumar, Souvik Kundu, Tatsuki Kuribayashi, Andrey Kutuzov

Matthieu Labeau, Wai Lam, Gerasimos Lampouras, Wuwei Lan, Yunshi Lan, Lukas Lange, Mirrella Lapata, Md Tahmid Rahman Laskar, Mark Last, Thai Le, Andrew Lee, Chia-Hsuan Lee, Dong-Ho Lee, En-Shiun Annie Lee, Hung-yi Lee, Hwanhee Lee, Jaeho Lee, Jay Yoon Lee, Mark G. Lee, Roy Ka-Wei Lee, SangKeun Lee, Seolhwa Lee, Tobias Lee, Shuo Lei, Wenqiang Lei, Heather Lent, Michael A. Lepori, Kristina Lerman, Piyawat Lertvittayakumjorn, Martha Lewis, Bangzheng Li, Bei Li, Chenliang Li, Chuanyi Li, Dongfang Li, Fei Li, Huayang Li, Irene Li, Jin Peng Li, Jing Li, Jingjing Li, Jiye Li, Junhui Li, Junyi Li, Lei Li, Li Li, Liangyou Li, Linjie Li, Linyang Li, Manling Li, Ming Li, Peifeng Li, Piji Li, Qian Li, Ru Li, Ruizhe Li, Shuai Li, Tao Li, Tianyi Li, Wenjie Li, Xiang Li, Xiaoli Li, Xin Li, Yafu Li, Yaliang Li, Yanyang Li, Yanzeng Li, Yaoyiran Li, Yinghui Li, Yinqiao Li, Yiyuan Li, Yongbin Li, Yu Li, Yuan-Fang Li, Zhenghua Li, Zhuang Li, Zixuan Li, Bin Liang, Hongru Liang, Paul Pu Liang, Xiaodan Liang, Lizi Liao, Jindřich Libovický, Constantine Lignos, Bill Y Lin, Hongfei Lin, Hongyu Lin, Hongzhan Lin, Hsien-chin Lin, Junyang Lin, Nankai Lin, Peiqin Lin, Shou-De Lin, Yankai Lin, Zhaojiang Lin, Bang Liu, Bingquan Liu, Dongqi Liu, Fenglin Liu, Guangliang Liu, Hao Liu, Hui Liu, Jiahao Liu, Jian Liu, Kang Liu, Ming Liu, Qi Liu, Qian Liu, Qianying Liu, Qun Liu, Tianyu Liu, Ting Liu, Tingwen Liu, Wei Liu, Xiao Liu, Xin Liu, Yang Janet Liu, Ye Liu, Yiqun Liu, Yixin Liu, Yong-Kang Liu, Zhenghao Liu, Zhiyuan Liu, Zoey Liu, Zuozhu Liu, Kare Livescu, Adam Lopez, Reza Lou, Charles Lovering, Hongyuan Lu, Wei Lu, Weiming Lu, Yaojie Lu, Li Lucy, Michal Lukasik, Junyu Luo, Weihua Luo, Xiao Luo, Ang Lv

Fenglong Ma, Jiaqi Ma, Jing Ma, Kaixin Ma, Mingyu Derek Ma, Xiaojuan Ma, Ziqiao Ma, Aman Madaan, Harish Tayyar Madabushi, Manuel Mager, Jean Maillard, Navonil Majumder, Chaitanya Malaviya, Shervin Malmasi, Valentin Malykh, Radhika Mamidi, Saab Mansour, Wenji Mao, Xin Mao, Yuning Mao, Zhuoyuan Mao, Kelly Marchisio, Marie-Catherine de Marneffe, Edison Marrese-Taylor, Giovanni Da San Martino, Sérgio Matos, Jonathan May, Joshua Maynez, Sahisnu Mazumder, Julian McAuley, Kathleen McKeown, Shikib Mehri, Clara Isabel Meister, Fandong Meng, Chunyan Miao, Fatemehsadat Mirehghallah, Kanishka Misra, Yasuhide Miura, Makoto Miwa, Fengran Mo, Lingbo Mo, Ashutosh Modi, Marie-Francine Moens, Hosein Mohebbi, Tansim Mohiuddin, Diego Molla, Syrielle Montariol, Nafise Sadat Moosavi, Roser Morante, Makoto Morishita, David R Mortensen, Marius Mosbach, Lili Mou, Aaron Mueller, Shamsuddeen Hassan Muhammad, Philippe Muller, Juan Pablo Munoz, Rudra Murthy

Masaaki Nagata, Ndapa Nakashole, Jinseok Nam, Blind Name, Jason Naradowsky, Vivi Nastase, Mir Tafseer Nayeem, Matteo Negri, Denis Newman-Griffis, Hwee Tou Ng, Dat Quoc Nguyen, Dong Nguyen, Hoang H Nguyen, Minh-Tien Nguyen, Ansong Ni, Shiwen Ni, Yixin Nie, Jan Niehues, Nafiseh Nikeghbal, Mitja Nikolaus, Takashi Ninomiya, Tong Niu, Farhad Nooralahzadeh, Pierre Nugues

Alice Oh, Naoaki Okazaki, Manabu Okumura, Arturo Oncevay-Marcos, Matan Orbach, Vicente Ordonez, Yohei Oseki, Nedjma Ousidhoum, Jessica Ouyang

Maria Leonor Pacheco, Vishakh Padmakumar, Sebastian Pado, Vardaan Pahuja, Liangming Pan, Jianhui Pang, Liang Pang, Isabel Papadimitriou, Sara Papi, Letitia Parcalabescu, Natalie Parde, Chan Young Park, Chanjun Park, Jong C. Park, Joonsuk Park, Kunwoo Park, Rebecca Passonneau, Panupong Pasupat, Viviana Patti, Siddharth Patwardhan, Ellie Pavlick, Jiaxin Pei, Hao Peng, Minlong Peng, Siyao Peng, Veronica Perez-Rosas, Fabio Petroni, Maxime Peyrard, Sandro Pezzelle, Matúš Pikuliak, Tiago Pimentel, Yuval Pinter, Irina Piontkovskaya, Benjamin Piwowarski, Bryan Allen Plummer, Simone Paolo Ponzetto, Hoifung Poon, Alexandros Potamianos, Shrimai Prabhu-

moye, Radityo Eko Prasajo, Daniel Preotiuc-Pietro, Emily Prud'hommeaux, Danish Pruthi, Yewen Pu

Chen Qian, Cheng Qian, Tiejun Qian, Chen Qiu, Han Qiu, Haoyi Qiu, Jieli Qiu, Xipeng Qiu, Lizhen Qu, Tingyu Qu, Xiaojun Quan

Vyas Raina, Dhananjay Ram, Alan Ramponi, Tharindu Ranasinghe, Surangika Ranathunga, Yanghui Rao, Hannah Rashkin, Shauli Ravfogel, Abhilasha Ravichander, Anton Razzhigaev, Traian Rebedea, Ines Rehbein, Marek Rei, Ehud Reiter, Navid Rekabsaz, HaoPeng Ren, Ruiyang Ren, Yi Ren, Zhaochun Ren, Leonardo Filipe Rodrigues Ribeiro, German Rigau, Shruti Rijhwani, Fabio Rinaldi, Anthony Rios, Oleg Rokhlenko, Julia Romberg, Salvatore Romeo, Benjamin Roth, Michael Roth, Joseph Le Roux, Alla Rozovskaya, Koustav Rudra, Ramon Ruiz-Dolz, Attapol Rutherford

Mrinmaya Sachan, Kenji Sagae, Swarnadeep Saha, Gözde Gül Sahin, Gaurav Sahu, Hassan Sajjad, Shailaja Keyur Sampat, Jivnesh Sandhan, Maarten Sap, Abulhair Saparov, Gabriele Sarti, Ryohei Sasano, Asad B. Sayeed, Viktor Schlegel, Michael Sejr Schlichtkrull, Fabian David Schmidt, Steven Schockaert, William Schuler, Raphael Schumann, H. Schwartz, Hinrich Schütze, João Sedoc, Indira Sen, Procheta Sen, Rico Sennrich, Maureen de Seyssel, Fei Sha, Lei Sha, Ankit Parag Shah, Guokan Shang, Jingbo Shang, Ori Shapira, Matthew Shardlow, Ehsan Shareghi, Vasu Sharma, Ravi Shekhar, Qinlan Shen, Xiaoyu Shen, Yongliang Shen, Tom Sherborne, Peng Shi, Tianze Shi, Weiyang Shi, Zhengliang Shi, Manish Shrivastava, Vered Shwartz, Jiasheng Si, Carina Silberer, Edwin Simpson, Maneesh K. Singh, Mayank Singh, Sunayana Sitaram, Konstantinos Skianis, Noah Smith, Mohammad Golam Sohrab, Aina Garí Soler, Haiyue Song, Linfeng Song, Wei Song, Xingyi Song, Yan Song, Yangqiu Song, Aitor Soroa, Gerasimos Spanakis, Richard Sproat, Vivek Srikumar, Balaji Vasan Srinivasan, Felix Stahlberg, Jacopo Staiano, Michal Stefanik, Elias Stengel-Eskin, Kevin Stowe, Marco Antonio Stranisci, Karl Stratos, Emma Strubell, Jannik Strötgen, Jinsong Su, Qi Su, Qinliang Su, Saku Sugawara, Alessandro Suglia, Elior Sulem, Aixin Sun, Fei Sun, Hao Sun, Jiao Sun, Jiashuo Sun, Jingyuan Sun, Kai Sun, Lichao Sun, Mingming Sun, Weiwei Sun, Yuan Sun, Mihai Surdeanu, György Szarvas

Santosh T.Y.S.S., Lynda Tamine Tamine, Minghuan Tan, Xingwei Tan, Zhaoxuan Tan, Zhixing Tan, Gongbo Tang, Hao Tang, Hongyin Tang, Raphael Tang, Siliang Tang, Tianyi Tang, Alberto Testoni, Christopher Thomas, Jesse Thomason, Yuanhe Tian, Zhiliang Tian, Antonio Toral, Amine Trabelsi, Quan Hung Tran, Thy Thy Tran, Richard Tzong-Han Tsai, Reut Tsarfaty, Yoshimasa Tsuruoka, Jingxuan Tu, Kewei Tu, Yunbin Tu, Martin Tutek

Ana Sabina Uban

Sowmya Vajjala, Marco Valentino, Jannis Vamvas, Neeraj Varshney, Olga Vechtomova, Alakananda Vempala, Pranav Narayanan Venkit, Suzan Verberne, Pat Verga, Karin Verspoor, Krishnapriya Vishnubhotla, Juraj Vladika, Rob Voigt, Soroush Vosoughi, Thuy Vu, Tu Vu, Yogarshi Vyas

Andreas Waldis, Byron C Wallace, Xiaojun Wan, Benyou Wang, Bin Wang, Bo Wang, Boxin Wang, Cheng Wang, Chengyu Wang, Dong Wang, Fei Wang, Guanchu Wang, Guoyin Wang, Haoyu Wang, Jiaan Wang, Jin Wang, Jindong Wang, Jingjing Wang, Keze Wang, Lei Wang, Liang Wang, Lingzhi Wang, Liwei Wang, PeiFeng Wang, Qingyun Wang, Rui Wang, Shijin Wang, Siyuan Wang, Wei Wang, Weiping Wang, Weiqi Wang, Wenjie Wang, Wenya Wang, Xi Wang, Xiaozhi Wang, Xin Wang, Xindi Wang, Xing Wang, Xinyu Wang, Xu Wang, Yan Wang, Yile Wang, Yu Wang, Yuxuan Wang, Zhaowei Wang, Zhiguang Wang, Zhongqing Wang, Zihao Wang, Zijian Wang, Leo Wanner, Alex Warstadt, Zeerak Waseem, Shinji Watanabe, Taro Watanabe, Tha-

rindu Cyril Weerasooriya, Wei Wei, Shira Wein, Leonie Weissweiler, Orion Weller, Haoyang Wen, Chris Wendler, Spencer Whitehead, Sarah Wiegrefe, Derry Wijaya, Genta Indra Winata, Kam-Fai Wong, Dustin Wright, Bingzhe Wu, Chien-Sheng Wu, Chuhan Wu, Fangzhao Wu, Hongqiu Wu, Lijun Wu, Qingyang Wu, Qiyu Wu, Te-Lin Wu, Xian Wu, Xiao-Ming Wu, Xiaobao Wu, Yiquan Wu, Yuanbin Wu, Yunfang Wu, Yuting Wu, Zhen Wu, Zhuofeng Wu

Ai Xi, Heming Xia, Peng Xia, Yang Xiang, Chaowei Xiao, Jing Xiao, Tong Xiao, Ziang Xiao, Qianqian Xie, Xin Xie, Zhouhang Xie, Bowen Xing, Chenyan Xiong, Deyi Xiong, Bo Xu, Chen Xu, Chunpu Xu, Frank F. Xu, Han Xu, Hongfei Xu, Lu Xu, Peng Xu, Qionghai Xu, Ruifeng Xu, Tong Xu, Yumo Xu, Zenglin Xu, Zhaozhao Xu, Nianwen Xue

Shuntaro Yada, Ikuya Yamada, An Yan, Hanqi Yan, Jianhao Yan, Jun Yan, Lingyong Yan, Rui Yan, Hitomi Yanaka, Chao-Han Huck Yang, Deqing Yang, Hao Yang, Jian Yang, Jie Yang, Linyi Yang, Min Yang, Pingjing Yang, Songlin Yang, Xianjun Yang, Yezhou Yang, Yinfei Yang, Zhao Yang, Zhenglu Yang, Zhiwei Yang, Zonglin Yang, Bingsheng Yao, Wenlin Yao, Yunzhi Yao, Ziyu Yao, Zonghai Yao, Michihiro Yasunaga, Mark Yatskar, Gregory Yauney, Fanghua Ye, Hai Ye, Jinyoung Yeo, Wen-tau Yih, Seid Muhie Yimam, Jian Yin, Qingyu Yin, Lin Yong, Kang Min Yoo, Naoki Yoshinaga, Koichiro Yoshino, Chenyu You, Haoxuan You, Dian Yu, Jianfei Yu, Lei Yu, Liang-Chih Yu, Tianshu Yu, Tong Yu, Wenhao Yu, Youngjae Yu, Yue Yu, Zhengtao Yu, David Yuan, Yifei Yuan, Zheng Yuan, Ma Yubo, Xiang Yue, Zhenrui Yue

Xingshan Zeng, Ziqian Zeng, Luke Zettlemoyer, Ce Zhang, Chen Zhang, Denghui Zhang, Hao Zhang, Haopeng Zhang, Jiajun Zhang, Jiaxin Zhang, Jieyu Zhang, Jing Zhang, Kai Zhang, Lei Zhang, Lemei Zhang, Li Zhang, Linfeng Zhang, Lu Zhang, Meishan Zhang, Min Zhang, Ming Zhang, Ningyu Zhang, Qi Zhang, Rongzhi Zhang, Rui Zhang, Ruiyi Zhang, Shaolei Zhang, Shichang Zhang, Shikun Zhang, Shiyue Zhang, Shuai Zhang, Shujian Zhang, Shuo Zhang, Taolin Zhang, Tong Zhang, Wei Zhang, Weinan Zhang, Wenqi Zhang, Wenxuan Zhang, Yan Zhang, Yang Zhang, Yi Zhang, Yifei Zhang, Yiming Zhang, Yong Zhang, Yuhao Zhang, Yuhui Zhang, Yunyi Zhang, Zeyu Zhang, Zhe Zhang, Zhuosheng Zhang, Chao Zhao, Chen Zhao, Chenye Zhao, Dongyan Zhao, Hai Zhao, Jieyu Zhao, Jun Zhao, Runcong Zhao, Tiejun Zhao, Wei Zhao, Xin Zhao, Xinran Zhao, Yang Yang Zhao, Zhixue Zhao, Bo Zheng, Chujie Zheng, Zaixiang Zheng, Victor Zhong, Ben Zhou, Bowen Zhou, Deyu Zhou, Guangyou Zhou, Hanzhang Zhou, Jiawei Zhou, Kaitlyn Zhou, Mingyang Zhou, Qingyu Zhou, Tianyi Zhou, Wangchunshu Zhou, Wenxuan Zhou, Yingbo Zhou, Yucheng Zhou, Hao Zhu, Qi Zhu, Qiaoming Zhu, Wanrong Zhu, Wenhao Zhu, Yilun Zhu, Zining Zhu, Honglei Zhuang, Terry Zhuo, Ayah Zirikly, Yftah Ziser, Shi Zong, Henry Peng Zou, Yicheng Zou, Arkaitz Zubiaga, Xinyu Zuo

## Reviewers

Ahmed Abbasi, Abdelrahman Abdallah, Mohamed Abdellatif, Hadi Abdine, Marzieh Abdolmaleki, Muhammad Abdul-Mageed, Amir Abdullah, Idris Abdulmumin, Giuseppe Abrami, Mitchell Abrams, Sharif Abuadbba, Arkadeep Acharya, Shantanu Acharya, Sharada Acharya, Emre Can Acikgoz, Samuel Ackerman, Eytan Adar, Tosin Adewumi, Suman Adhya, Yossi Adi, Dyah Adila, Muhammad Farid Adilazuarda, Helena Gomez Adorno, Haithem Afli, Tazin Afrin, Aishwarya Agarwal, Amit Agarwal, Ankush Agarwal, Ishika Agarwal, Mayank Agarwal, Saurabh Agarwal, Sumit Agarwal, Utkarsh Agarwal, Ameya Agaskar, Charu C. Aggarwal, Divyanshu Aggarwal, Karan Aggarwal, Ehsan Aghazadeh, Mohamed Aghzal, Eneko Agirre, Ayush Agrawal, Samarth Agrawal, Hafsa Ahmad, Zishan Ahmad, Arash Ahmadian, Nesreen Ahmed, Shafiuddin Rehan Ahmed, Zeeshan Ahmed, Jaewoo Ahn, Sumyeong Ahn, Wonhyuk Ahn, Sanchit Ahuja, Lin Ai, Qingyao Ai, Wei Ai, Taichi Aida, Akiko Aizawa, Tunde Oluwaseyi Ajayi, Pritom Saha Akash,



Benjamin Akera, Mubashara Akhtar, Kosuke Akimoto, Chandan Akiti, Berfin Aktas, Syeda Nahida Akter, Hend Al-Khalifa, Hadeel Al-Negheimish, Nora Al-Twairesh, Badr AlKhamissi, Hilal AlQuabeh, Jesujoba Oluwadara Alabi, Ozge Alacam, Ahmed Alajrami, Belen Alastruey, Lisa Alazraki, Alon Albalak, Itziar Aldabe, Hanan Aldarmaki, Georgios Alexandridis, Enrique Alfonseca, Robin Jonathan Algayres, Bashar Alhafni, Raquel G. Alhama, Felermينو D. M. A. Ali, Iqra Ali, Mehdi Ali, Mubashir Ali, Ilseyar Alimova, Pegah Alipoormolabashi, Dan Alistarh, Tamer Alkhoul, Ali Allaith, Alexandre Allauzen, Nada Almarwani, Deema Alnuhait, Hani Alomari, Iñigo Alonso, Kenneth Alperin, Sophia Althammer, Duarte Miguel Alves, Kshitij Alwadhi, Fakhraddin Alwajih, Zaid Alyafeai, Norah A. Alzahrani, Mahsa Amani, Alfonso Amayuelas, Jacopo Amidei, Enrique Amigo, Evelin Amorim, Samuel Joseph Amouyal, Haozhe An, Jisun An, Kaikai An, Na Min An, Shengnan An, Wenbin An, Konjengbam Anand, Suraj Anand, Amith Ananthram, Nicholas Andrews, Ion Androustopoulos, Rico Angell, Krasimir Angelov, Simon D Angus, Tatiana Anikina, Akash Anil, Ahmad Mustafa Anis, Murali Annavaram, Miriam Anschütz, Wang Ante, Wissam Antoun, David Anugraha, Gopala Anumanchipalli, Rao Muhammad Anwer, Junyi Ao, Liu Aofan, Yoichi Aoki, Tatsuya Aoyama, Alessio Palmero Aprosio, Mario Ezra Aragón, Erik Arakelyan, Pedro Henrique Luz de Araujo, Jerónimo Arenas-García, Esteban Garces Arias, Samee Arif, Sercan O Arik, Nurul Fajrin Ariyani, Konstantine Arkoudas, Jordi Armengol-Estapé, Cristian Berrio Aroca, Akshatha Arodi, Akhil Arora, Arnav Arora, Aryaman Arora, Gaurav Arora, David Arps, Katya Artemova, Mikel Artetxe, Keren Artiaga, Mohammad Arvan, Masaki Asada, Yuki Asano, Ali Asgarov, Nicholas Asher, Hadi Askari, Dennis Assenmacher, Ira Assent, Md. Atabuzzaman, Duygu Ataman, Thushari Atapattu, Awais Athar, Ben Athiwaratkun, Yash Kumar Atri, Omar Attia, Aitziber Atutxa, Berk Atıl, Javier Aula-Blasco, Mikko Aulamo, Nathalie Aussenac-Gilles, P. V. S. Avinesh, Andrei Marius Avram, Kumar Ayush, Hammad Ayyubi, Asif Azad, Amos Azaria, Abdul Hameed Azeemi, Israel Abebe Azime, Dilshod Azizov

Yukino Baba, Nikolay Babakov, Petr Babkin, Stephen Bach, Michael Backes, Sher Badshah, Kyunghoon Bae, Sanghwan Bae, Suyoung Bae, Jinheon Baek, Joachim De Baer, Bart Baesens, Roxanne El Baff, Niyati Bafna, Mahdiah Baghshah, Seyed Ali Bahrainian, Haoli Bai, Huajun Bai, Jing Bai, Jun Bai, Rui Bai, Xiaoying Bai, Yu Bai, Yuelin Bai, Yushi Bai, Yuzhuo Bai, Ziyi Bai, Pra-soon Bajpai, Aleksandra Bakalova, George Arthur Baker, Yavuz Faruk Bakman, Bal Krishna Bal, Murali Balakrishnan, Oana Denisa Balalau, Nihal Balani, Aruna Balasubramanian, Rishab Balasubramanian, Nishant Balepur, Nicolas Ballier, Simone Balloccu, Lucas Bandarkar, Dibyanayan Bandyopadhyay, Sambaran Bandyopadhyay, Somnath Banerjee, Yejin Bang, Naman Bansal, Calvin Bao, Guangsheng Bao, Jianzhu Bao, Junwei Bao, Qiming Bao, Rong Bao, Runxue Bao, Siqi Bao, Yujia Bao, Kfir Bar, Libby Barak, Lorenzo Baraldi, Alexander Baranov, Edoardo Barba, M Saiful Bari, Jeremy Barnes, Ander Barrena, James Barry, Martijn Bartelds, Madhusudan Basak, Advik Raj Basani, Pierpaolo Basile, Elisa Bassignana, Osbert Bastani, Kinjal Basu, Samyadeep Basu, Riza Batista-Navarro, Anil Batra, Alessia Battisti, David Bau, Irene Baucells, John Bauer, Lisa Bauer, Tim Baumgärtner, Connor Baumler, Rachel Bawden, Deniz Bayazit, Adrien Bazoge, Nicholas Beauchamp, Srikantha J. Bedathur, Steven Bedrick, Barend Beekhuizen, Peter Anthony Beerel, Parishad BehnamGhader, Andrei Stefan Bejgu, Gemma Bel-Enguix, Tadesse Destaw Belay, Peter Belcak, Frederico Belcavello, Lamia Belguith, Peter Bell, Tony Belpaeme, Boualem Benatallah, Michael Bendersky, Luca Benedetto, Himanshu Beniwal, Ivana Benova, Luisa Benti-vogli, Christian Bentz, Gábor Berend, Timothée Bernard, William Berrios, Dario Bertero, Amanda Bertsch, Katrien Beuls, Sourav Sanjukta Bhabesh, Milan Bhan, Abhishek Bhandari, Manik Bhandari, Manasa Bharadwaj, Samarth Bhargav, Riyaz Ahmad Bhat, Gagan Bhatia, Shaily Bhatt, Amrita Bhattacharjee, Arnab Bhattacharya, Indrajit Bhattacharya, Sourangshu Bhattacharya, Uttaran Bhattacharya, Pushpak Bhattacharyya, Ankita Bhaumik, Nikhil Bhendawade, Ayan Kumar Bhowmick, Rajarshi Bhowmik, Amran Bhuiyan, Nikita Bhutani, Bin Bi, Guanqun Bi, Keping Bi, Ziqian Bi, Przemyslaw Biecek, Maria Bielikova, Laura Biester, Iman Munire Bilal, Yi Bin, Saman Bookaghazadeh, Shankar Biradar, Arianna Bisazza, Tegawendé F. Bissyandé, Danielle Bitterman,

Johanna Björklund, Frederic Blain, Verena Blaschke, Sravan Babu Bodapati, Purna Srivatsa Bogavarapu, Joanne Boisson, Marcelly Zanon Boito, Ondrej Bojar, Brendon Boldt, Marco Bombieri, Chiara Di Bonaventura, Philipp Borchert, Łukasz Borchmann, Georgeta Bordea, Claudia Borg, Beatriz Borges, Zeinab Borhanifard, Hsuvas Borkakoty, Matteo Bortoletto, Cristina Bosco, Dibalay Bose, Sombit Bose, Wen Bosi, Stefan Bott, Dani Bouch, Sabri Boughorbel, Andrey Bout, Leonid Boytsov, Danielle Bragg, Maharaj Brahma, Faeze Brahman, Ana Brassard, Daniel Braun, Felipe Bravo-Marquez, Cynthia Breazeal, Eleftheria Briakou, Jannik Brinkmann, Ted Briscoe, Tomás Vergara Browne, Elia Bruni, Alessio Brutti, Jiajun Bu, Mengyu Bu, Weerayut Buaphet, Ana-Maria Bucur, Arun Buduru, Kyle Robert Buettner, Nam Khac-Hoai Bui, Trung Bui, Tony Santiago Montes Buitrago, Andreas Bulling, Razvan Bunescu, Laurie Burchell, Sophie Burkhardt, Evgeny Burnaev, Andrea Burns, Victor Bursztyn, Hendrik Buschmeier, Miriam Butt, Sabur Butt, Evgenii V Bykovets, Bill Byrne, SungJoo Byun, Necva Bölücü

Isabel Cachola, Luca Cagliero, Samuel Cahyawijaya, Erica Cai, Fengyu Cai, Guohao Cai, Hengrui Cai, Hengyi Cai, Kunlin Cai, Li Cai, Min Cai, Pengshan Cai, Ruichu Cai, Shu-Bin Cai, Xiaoyan Cai, Xunliang Cai, Yi Cai, Yujun Cai, Yuliang Cai, Yunfeng Cai, Zefan Cai, Zhenyang Cai, Agostina Calabrese, Ioan Calapodescu, Aylin Caliskan, Chris Callison-Burch, Eduardo Calò, Nguyen Cam-Tu, Jose Camacho-Collados, Erik Cambria, Stefano Campese, Daniel F Campos, Duy-Cat Can, Marc Canby, Nicola Cancedda, Bin Cao, Bochuan Cao, Bowen Cao, Boxi Cao, Chengyu Cao, Cong Cao, Congfeng Cao, Hailong Cao, Hannan Cao, Hongye Cao, Juan Cao, Juntai Cao, Kris Cao, Lang Cao, Qi Cao, Qian Cao, Ruisheng Cao, Shulin Cao, Shuyang Cao, Siqi Cao, Xuezhi Cao, Yanshuai Cao, Yixuan Cao, Yu Cao, Yuan Cao, Yuanpu Cao, Yuhan Cao, Yupeng Cao, Yushi Cao, Zhao Cao, Zhenbiao Cao, Zhiyu Cao, Cornelia Caragea, Ronald Cardenas, Rémi Cardon, Henrique Lopes Cardoso, Mark Carman, Danilo Carvalho, Edresson Casanova, Paola Cascante-Bonilla, Tommaso Caselli, Silvia Casola, Pierluigi Cassotti, Ignacio Castro, Alberto Castro-Hernandez, Camilla Casula, Isaac Rayburn Caswell, Bryan Catanzaro, Paulo Rodrigo Cavalin, Francesco Cazzaro, Jan Cegin, Yukuo Cen, Aman Chadha, Dong-Kyu Chae, Joyce Chai, Linzheng Chai, Mingxu Chai, Wenhao Chai, Yekun Chai, Aunabil Chakma, Soumen Chakrabarti, Abhijit Chakraborty, Souradip Chakraborty, Sunandan Chakraborty, Gaël de Chalendar, Fateme Hashemi Chaleshtori, Alvin Chan, Chee Seng Chan, Chimin Chan, David Chan, Hou Pong Chan, Luo Qi Chan, Sophia Chan, Navin Chandak, Mohit Chandra, Vikas Chandra, Aofei Chang, Buru Chang, Du-Seong Chang, Heng Chang, Jonathan P. Chang, Kai-Wei Chang, Ma Chang, Serina Chang, Simyung Chang, Tyler A Chang, Yapei Chang, Zhiyuan Chang, Yang Chao, Lucas Georges Gabriel Charpentier, Anwoy Chatterjee, Usashi Chatterjee, Akshay Chaturvedi, Sagar Chaturvedi, Mainak Chaudhury, Geeticka Chauhan, Kushal Chawla, Nitesh Chawla, Andong Chen, Anni Chen, Anthony Chen, Bei Chen, Beiduo Chen, Ben Chen, Bin Chen, Bingsen Chen, Bo Chen, Bowen Chen, Boxing Chen, C. L. Philip Chen, Canyu Chen, Catherine Chen, Cen Chen, Chacha Chen, Changyu Chen, Charlie Chen, Chen Chen, Chih Yao Chen, Chunchun Chen, Congliang Chen, Danlu Chen, Daoyuan Chen, DeLong Chen, Dongdong Chen, Dongping Chen, Dubing Chen, Enhong Chen, Feng Chen, Gang Chen, Guiming Hardy Chen, Guizhen Chen, Guoxin Chen, Haifeng Chen, Haihua Chen, Hang Chen, Hanjie Chen, Hannah Chen, Hanzhu Chen, Hao Chen, Hong Chen, Hong-You Chen, Hongjun Chen, Hongzhan Chen, Huajun Chen, Huimin Chen, Huiyao Chen, Jiaju Chen, Jiajun Chen, Jian Chen, Jiaqi Chen, Jiawei Chen, Jiaxiang Chen, Jinghong Chen, Jiyu Chen, John Chen, Junxi Chen, Junying Chen, Junzhe Chen, Kai Chen, Kaiyuan Chen, Kang Chen, Kangjie Chen, Ke Chen, Kedi Chen, Kejiang Chen, Kuang-Ming Chen, Lihu Chen, Ling Chen, Liyang Chen, Lizhong Chen, Longze Chen, Luoyu Chen, Lyuhao Chen, Meiqi Chen, Meng Chen, Nuo Chen, Pei Chen, Qi Chen, Qian Chen, Qiang Chen, Qiguang Chen, Qin Chen, Qingcai Chen, Qingyu Chen, Qizhou Chen, Qun Chen, Ruijun Chen, Ruirui Chen, Shan Chen, Shijie Chen, Shisong Chen, Shiwei Chen, Shuang Chen, Shunian Chen, Shuo Chen, Si Chen, Siheng Chen, Simin Chen, Sishuo Chen, Tao Chen, Tianxiang Chen, Tong Chen, Wei Chen, Weize Chen, Wenqing Chen, Wenting Chen, Wenyu Chen, Xi Chen, Xiang Chen, Xiaojun Chen, Xie

Chen, Xinyue Chen, Xu Chen, Xuanang Chen, Xun Chen, Yanfei Chen, Yankai Chen, Yanquan Chen, Yanran Chen, Yi-Pei Chen, Yige Chen, Yilong Chen, Yimeng Chen, Yin-Jen Chen, Yingfa Chen, Yiqun Chen, Yiran Chen, Yirong Chen, Yiwen Chen, Yixin Chen, Yiyi Chen, Yongrui Chen, Yubo Chen, Yuen Chen, Yufeng Chen, Yuhan Chen, Yulin Chen, Yun Chen, Yuxuan Chen, Yuyan Chen, Zeming Chen, Zhanpeng Chen, Zhaoling Chen, Zhaorun Chen, Zhe Chen, Zhengyu Chen, Zhi-Yuan Chen, Zhihong Chen, Zhipeng Chen, Zhongwu Chen, Zhouxi Chen, Zhu-min Chen, Zhuo Chen, Zhuowei Chen, Zhuoxiao Chen, Ziyang Chen, Zongxiong Chen, Zulong Chen, Daixuan Cheng, Dawei Cheng, Gang Cheng, Hong Cheng, Jiale Cheng, Jian Cheng, Jianpeng Cheng, Jiayang Cheng, Jingwei Cheng, Julius Cheng, Kanzhi Cheng, Liang Cheng, Minhao Cheng, Myra Cheng, Pengxiang Cheng, Pengzhou Cheng, Qiao Cheng, Qinyuan Cheng, Shanbo Cheng, Shaohuan Cheng, Sijie Cheng, Siyuan Cheng, Wei Cheng, Xiaoxue Cheng, Xize Cheng, Xuxin Cheng, Yao Cheng, Yi Cheng, Ying Cheng, Yuan Cheng, Zhi-Qi Cheng, Zihui Cheng, Liu Chengwu, Yun-Gyung Cheong, Colin Cherry, Emmanuele Chersoni, Jackie Ck Cheung, William Cheung, Rendi Chevi, Jie Chi, Yew Ken Chia, Cheng-Han Chiang, David Chiang, Ting-Rui Chiang, Jenny Chim, Si-Chi Chin, Elena Chistova, Pranjal A Chitale, Yu Ying Chiu, Odysseas S. Chlapanis, Eng Siong Chng, Hyowon Cho, Hyundong Justin Cho, Hyunsouk Cho, Ikhyun Cho, Ji Yong Cho, Kyunghyun Cho, Won Ik Cho, Young Min Cho, Donghee Choi, Eunkyung Choi, Eunseong Choi, Jaesik Choi, Jihun Choi, Joon-Young Choi, Juhwan Choi, Jungwook Choi, Kwanghee Choi, Seungtaek Choi, YongSuk Choi, Yubin Choi, Shamil Chollampatt, Christopher A. Choquette-Choo, Mukund Choudhary, Munmun De Choudhury, Sagnik Ray Choudhury, Yinlam Chow, Arijit Ghosh Chowdhury, Sankalan Pal Chowdhury, Minh Duc Chu, Xiaowen Chu, Xu Chu, Yun-Wei Chu, Zheng Chu, Tat-seng Chua, Lv Chuancheng, Yu-Neng Chuang, Yun-Shiuan Chuang, Yung-Sung Chuang, Chanyoung Chung, Ho-Lam Chung, Jin-Woo Chung, Jiwan Chung, John Joon Young Chung, Tagyoung Chung, Lu Chunlin, Alessandra Teresa Cignarella, Philipp Cimiano, Christian Clark, Thomas Hikaru Clark, Chloé Clavel, Stéphane Clinchant, Amir David Nissan Cohen, Andrew Cohen, Roi Cohen, Vanya Cohen, Pierre Colombo, Tianshuo Cong, Xin Cong, Mathieu Constant, Owen Cook, Paul Cook, Rodolfo Corona, Ander Corral, Adrian Cosma, Marta R. Costa-jussà, Mathias Creutz, Danilo Croce, Jan Christian Blaise Cruz, Ganqu Cui, Jiajia Cui, Lizhen Cui, Peng Cui, Shaobo Cui, Shiyao Cui, Shuguang Cui, Wanyun Cui, Wendi Cui, Yuanning Cui, Washington Cunha, Alba Cercas Curry, Paula Czarnowska, Bingfeng chen, Erion Çano, Buse Çarık

Mahavir Dabas, Nico Daheim, Deborah A. Dahl, Daniel Dahlmeier, Bin Dai, Chuyi Dai, Gao-le Dai, Hongliang Dai, Huangyu Dai, Quanyu Dai, Xiang Dai, Xinyu Dai, Yong Dai, Parag Pravin Dakle, Dwip Dalal, David Dale, Majid Daliri, Jeff Dalton, Elizabeth M. Daly, Suyash Damle, Greta Damo, Soham Dan, Yuhao Dan, Ankit Dangi, Mai Hoang Dao, Zhang Daoan, Trevor Darrell, Kareem Mohamed Darwish, Anirban Das, Dipankar Das, Sarkar Snigdha Sarathi Das, Souvik Das, Mihai Dascalu, Saloni Dash, Debtanu Datta, Sam Davidson, Arka Daw, Hillary Dawkins, Jay DeYoung, Nicholas Deas, Stefania Degaetano-Ortlieb, Hiroyuki Deguchi, Derrick Goh Xin Deik, Hanieh Deilamsalehy, Louise Deleger, David Demeter, Chunyuan Deng, Guanzhi Deng, Hexuan Deng, Jiawen Deng, Jieren Deng, Jingcheng Deng, Naihao Deng, Shuiguang Deng, Weiwei Deng, Wentao Deng, Yue Deng, Zhenyun Deng, Zheyue Deng, Zhi-Hong Deng, Zhijie Deng, Zhongfen Deng, Sourabh Dattatray Deoghare, Leon Derczynski, Jan Milan Deriu, Franck DERNONCOURT, Aniket Deroy, Akshar Prabhu Desai, Tejash Desai, Ameet S Deshpande, Anushka Deshpande, Vijeta Deshpande, Nina Dethlefs, Chaitanya Devaguptapu, Prajit Dhar, Ruchira Dhar, Bhuwan Dhingra, Kaustubh Dhole, Shehzaad Zuzar Dhuliawala, Barbara DiEugenio, Enmao Diao, Xiaolei Diao, Xingjian Diao, Gaël Dias, Fernando Diaz, Harshita Diddee, Adji Bouso Dieng, Stefan Dietze, Dimitar Dimitrov, Dimitar Iliyanov Dimitrov, Emily Dinan, Jacob Dineen, Bolin Ding, Chenchen Ding, Chuanghao Ding, Dujian Ding, Hanxing Ding, Jian Ding, Keyan Ding, Nai Ding, Ning Ding, Peng Ding, Ruixue Ding, Shuoyang Ding, Tong Ding, Wentao Ding, Wenxuan Ding, Xiao Ding, Yi Ding, Yifan Ding, Yifeng Ding, Yifu Ding, Yihao Ding, Yuyang Ding, Zifeng Ding,

Song Dingjie, Ly Dinh, Tanvi Dinkar, Shubhashis Roy Dipta, Ajay Divakaran, Anuj Diwan, Lucas Dixon, Amirbek Djanibekov, Heejin Do, Jaeyoung Do, Phong Nguyen-Thuan Do, Quyet V. Do, Khoa Doan, Sumanth Doddapaneni, Rama Doddipatla, Pierre Dognin, Kosuke Doi, Carlotta Domeniconi, Tobias Domhan, Bin Dong, Chengyu Dong, Guanting Dong, Haoyu Dong, Jianfeng Dong, Junnan Dong, Ming Dong, Mingwen Dong, Peijie Dong, Rui Dong, Wenchao Dong, Xiangjue Dong, Xin Luna Dong, Yushun Dong, Zhe Dong, Zhengyuan Dong, Zican Dong, Vardhan Dongre, Ehsan Doostmohammadi, Emmanuel Dorley, Meet Doshi, Bonaventure F. P. Dossou, Longxu Dou, Qingyun Dou, Yao Dou, Mohammad Javad Dousti, Mauro Dragoni, Mark Dras, Mark Dredze, Felix Drinkall, Daniel Dsouza, Haowei Du, Jiangshu Du, Jinhua Du, Kevin Du, Lun Du, Minxin Du, Nan Du, Pan Du, Weiyu Du, Wenyu Du, Xiaoyong Du, Yongkang Du, Yypei Du, Zhihao Du, Hanyu Duan, Jinhao Duan, Jiuding Duan, Juanyong Duan, Lei Duan, Lixin Duan, Siyuan Duan, Sufeng Duan, Xinyu Duan, Xufeng Duan, Zenghao Duan, Deep Narayan Dubey, Haim Dubossarsky, Richard Dufour, Liam Dugan, Kevin Duh, Ambedkar Dukkipati, Lavinia Dunagan, Jocelyn Dunstan, Emmanuel Dupoux, Benjamin Van Durme, Ritam Dutt, Subhabrata Dutta, Sujun Dutta, Jennifer Dy, Nils Dycke, Pius von Däniken, Esra Dönmez

Haihong E, Seth Ebner, Jessica Maria Echterhoff, Ghebriout Mohamed Imed Eddine, Lilach Eden, Tobias Eder, Aleksandra Edwards, Paul Van Eecke, Steffen Eger, Koji Eguchi, Yo Ehara, Alon Eirew, Jason Eisner, Mai ElSherief, Mohamed Elaraby, Hoda Eldardiry, Mohamed Elgaar, Mohamed Elhoseiny, Tilia Ellendorff, AbdelRahim Elmadany, Tamer Elsayed, Yahya Emara, Chris Emmerly, Sugyeong Eo, Elena V. Epure, Pierre Erbacher, Erkut Erdem, Justus-Jonas Christoph Erker, Marieke van Erp, Hugo Jair Escalante, Suilan Estevez-Velarde, Thierry Etchegoyhen, Julen Etxaniz, Kilian Evang, Keelan Evanini, Gautier Evennou, Brendan Raffel Evers, Chandra Kiran Reddy Evuru, Matan Eyal

Marzieh Fadaee, Ekaterina Fadeeva, Md Fahim, Fahim Faisal, Júlia Falcão, Neele Falk, Caoyun Fan, Chenrui Fan, Cunhang Fan, Jiabin Fan, Jing Fan, Ming Fan, Shengda Fan, Wei Fan, Xiaoran Fan, Xiran Fan, Yongqi Fan, Yue Fan, Zhaoxin Fan, Zhiyuan Fan, Anjie Fang, Chao Fang, Chunrong Fang, Fang Fang, Feiteng Fang, Haishuo Fang, Jinyuan Fang, Ming Fang, Minghong Fang, Minghui Fang, Qingkai Fang, Tao Fang, Tianqing Fang, Xinyu Fang, Yi Fang, Yuan Fang, Reza Farahbakhsh, António Farinhas, Pedro Faustini, Barah Fazili, Arnisa Fazla, Pooyan Fazli, Maria Fedorova, Xiang Fei, Yu Fei, Zhaoye Fei, Zhiwei Fei, Bao Feilong, Nils Feldhus, Sergey Feldman, Andrew Zhuoer Feng, Aosong Feng, Chong Feng, Davy Feng, Fangxiang Feng, Fuli Feng, Guanchao Feng, Guhao Feng, Hao Feng, Junlan Feng, Lei Feng, Ling Feng, Rui Feng, Shutong Feng, Steven Y. Feng, Tao Feng, Tian Feng, Xincan Feng, Yang Feng, Yansong Feng, Yi Feng, Yu Feng, Yubo Feng, Yujie Feng, Yunlong Feng, Zhiyong Feng, Zihao Feng, Zijian Feng, Pan Fengjun, Alena Fenogenova, Teddy Ferdinan, Patrick Fernandes, Nigel Fernandez, Aloka Fernando, Basura Fernando, Raquel Fernández, Alberte Fernández-Castro, Elisa Ferracane, Eric Le Ferrand, Emilio Ferrara, Thomas Palmeira Ferraz, Elisabetta Fersini, Sheridan Feucht, Blanca Calvo Figueras, Georgios Filandrianos, Elena Filatova, Jillian Fisher, Marco Fisichella, Jack FitzGerald, Pieter Fivez, Lorenzo Jaime Yu Flores, Achille Fokoue, Marcio Fonseca, Negar Foroutan, Stella Frank, Thomas François, Diego Frassinelli, Dayne Freitag, Jibril Frej, Simona Frenda, Giacomo Frisoni, Michael Fromm, Dayuan Fu, Harvey Yiyun Fu, Tingchen Fu, YiFu Fu, Yicheng Fu, Yiwei Fu, Yu Fu, Yun Fu, Yuzhuo Fu, Zichuan Fu, Dennis Fucci, Yoshinari Fujinuma, Atsushi Fujita, Yasuhiro Fujiwara, Richard Futrell

Kun Gai, Varun Kumar Reddy Gajjala, Jay Gala, Dimitris Galanis, Marcos Martínez Galindo, Lukas Paul Achatius Galke, Martina Galletti, Jack Gallifant, Aram Galstyan, Pablo Gamallo, Andrew Gambardella, Xiaoying Gan, Yanglei Gan, Yujian Gan, Sriram Ganapathy, Adithya V Ganesan, Balaji Ganesan, Rashmi Gangadharaiah, Varun Prashant Gangal, Debasis Ganguly, William Gantt, Bofei Gao, Changjiang Gao, Chao Gao, Chaochen Gao, Chen Gao, Chongming Gao, Chon-

gyang Gao, Chufan Gao, Cuiyun Gao, Dehong Gao, Feiyu Gao, Ge Gao, Guangze Gao, Hang Gao, Hongcheng Gao, Hui Gao, Jiahui Gao, Jianfeng Gao, Jing Gao, Jinglong Gao, Jingsheng Gao, Jinhua Gao, Jinyang Gao, Junqi Gao, Lang Gao, Mingqi Gao, Muhan Gao, Neng Gao, Qiyue Gao, Rena Wei Gao, Shen Gao, Shengxiang Gao, Silin Gao, Wei Gao, Xiaoxue Gao, Yanjun Gao, Yingqiang Gao, Zhujin Gao, Utpal Garain, Iker García-Ferrero, Claire Gardent, Nikesh Garera, Ankur Garg, Dinesh Garg, Siddhant Garg, Vaibhav Garg, Aparna Garimella, Joseph Gatto, Himanshu Gauba, Vagrant Gautam, Daniil Gavrilov, Jiaxin Ge, Mengshi Ge, Xiou Ge, Yubin Ge, Zongyuan Ge, Gregor Geigle, Spandana Gella, Debela Tesfaye Gemechu, Josef van Genabith, Saibo Geng, Yuxia Geng, Ariel Gera, Michael Gertz, Walter Gerych, Omid Ghahroodi, Soumya Suvra Ghosal, Akash Ghosh, Kripabandhu Ghosh, Saptarshi Ghosh, Sayan Ghosh, Sayontan Ghosh, Shaona Ghosh, Sreyan Ghosh, Sucheta Ghosh, George Giannakopoulos, Fosca Giannotti, C. Lee Giles, Luca Gioffrè, Raja Giryes, Adrià de Gispert, Fausto Giunchiglia, Michael Glass, Oren Glickman, Kristina Gligoric, Aaron Gluck, Satya Sai Srinath Namburi Gnvv, Ameya Godbole, Randy Goebel, Anmol Goel, Arnav Goel, Iakes Goenaga, Sim Kuan Goh, Usman Gohar, Darina Gold, Adele E. Goldberg, Yoav Goldberg, Jonas Golde, Tomas Goldsack, Tom Goldstein, Sujatha Das Gollapalli, Michal Golovanevsky, Jose Manuel Gomez-Perez, Marcos A Goncalves, Elizaveta Goncharova, Chen Gong, Hailei Gong, Jie Gong, Kaixiong Gong, Ming Gong, Neil Gong, Shansan Gong, Yeyun Gong, Yunye Gong, Zhiguo Gong, Ziwei Gong, Joseph E. Gonzalez, Cesar Gonzalez-Gutierrez, Julio Gonzalo, Noah Goodman, Deepak Gopinath, Dhiman Goswami, Sujata Goswami, Daniela Gottesman, Fabrizio Gotti, Boyu Gou, Yunhao Gou, Pierre-Antoine Gourraud, Antoine Gourru, Venkata Subrahmanyam Govindarajan, Roksana Goworek, Agam Goyal, Palash Goyal, Pushap Goyal, Vikram Goyal, Yash Goyal, Matthias Grabmair, Filip Graliński, Christan Grant, Tommaso Green, Rachel Greenstadt, Matt Grenander, Georg Groh, Rynaa Grover, Adam Grycner, Hengrui Gu, Jinghang Gu, Naibin Gu, Tiancheng Gu, Tianle Gu, Tianlong Gu, Xiaodong Gu, Xiaotao Gu, Yi Gu, Yuxuan Gu, Zhouhong Gu, Eleonora Gualdoni, Jian Guan, Nan Guan, Quanlong Guan, Shuhao Guan, Weili Guan, Xin Guan, Xinyan Guan, Xinze Guan, Ziyu Guan, Li Hong Guang, Peter Guang, Martin Gubri, Marco Guerini, Nuno M Guerreiro, Arjun Guha, Shouvik Kumar Guha, Liane Guillou, Sumit Gulwani, Varun Gumma, Satya Kesav Gundabathula, Geyang Guo, Hangyu Guo, Hongcheng Guo, Huifeng Guo, Jiang Guo, Jiaxin Guo, Junjun Guo, Lingbing Guo, Meiqi Guo, Pei-Fu Guo, Ping Guo, Qingyan Guo, Qipeng Guo, Quanjiang Guo, Rui Guo, Ruohao Guo, Shangwei Guo, Shaoru Guo, Shoutao Guo, Siyuan Guo, Taicheng Guo, Xiaobo Guo, Xinnan Guo, Xu Guo, Yuhang Guo, Zhicheng Guo, Zhiqiang Guo, Akash Gupta, Akshat Gupta, Ankita Gupta, Ashim Gupta, Lavanya Gupta, Priyanshu Gupta, Rahul Gupta, Sameer Gupta, Shivanshu Gupta, Sonam Gupta, Iryna Gurevych, Sireesh Gururaja, Ximena Gutierrez-Vasques, Alexander Gutkin, Daehoon Gwak, Ingo Gühring

Le An Ha, Hakim Hacid, Shohreh Haddadan, Nikita Haduong, Lovisa Hagström, Joonghyuk Hahn, Thomas Hain, Mao Haitao, Jan Hajič, Sheikh Azizul Hakim, Zaber Ibn Abdul Hakim, Ahmad Dawar Hakimi, Sherzod Hakimov, Rajarshi Haldar, Scott A. Hale, Karina H Halevy, Coleman Haley, Patrick Haller, Skyler Hallinan, Injy Hamed, Caren Han, Cheng Han, Chengcheng Han, Dou Shi Han, Haoyu Han, Jiawei Han, Jinyoung Han, Jiuzhou Han, Jizhong Han, Kanyao Han, Kelvin Han, Meng Han, Namgi Han, Seungju Han, Simeng Han, William Han, Xiaotian Han, Xiongwei Han, Xu Han, Xudong Han, Xue Han, Yo-Sub Han, Kunal Handa, Adeep Hande, Zhang Hanlin, Chuzhan Hao, Hailin Hao, Jianye Hao, Liu Hao, Tianyong Hao, Xiaoyang Hao, Yaru Hao, Yongchang Hao, Zhifeng Hao, Fang Haonan, Saiful Haq, Rejwanul Haque, Sanda Harabagiu, Momchil Hardalov, Christian Hardmeier, Fabrice Y Harel-Canada, Stefan Harmeling, Ian Harris, Joshua Hartshorne, Thomas Hartvigsen, Md Kamrul Hasan, Md Rakibul Hasan, Md. Arid Hasan, Mohammad Rashedul Hasan, Maram Hasanain, Taku Hasegawa, Mark A. Hasegawa-Johnson, Farnoosh Hashemi, Faegheh Hasibi, Md. Mahadi Hassan, Michael Hassid, Avinatan Hassidim, Hans Ole Hatzel, Bradley Hauer, Shreya Havaladar, Majd Hawasly, Hiroaki Hayashi, Katsuhiko Hayashi, Kazuki Hayashi, Amir Hazem, Bingsheng He, Bingxiang He, Bolei He, Bolei He, Bolei He, Chaoqun

He, Conghui He, Di He, Estrid He, Guoxiu He, Hangfeng He, Hao He, Haorui He, Jianfeng He, Jiayi He, Junda He, Kai He, Kaiyu He, Kun He, Liang He, Liyang He, Mingqian He, Pengfei He, Qianyu He, Qing He, Ruifang He, Shwai He, Tingting He, Wei He, Xiangjian He, Xiangnan He, Xingwei He, Xuming He, Yancheng He, Yifei He, Yongquan He, Yueru He, Yun He, Yuxiong He, Zexue He, Zhankui He, Zhaofeng He, Zhitao He, Zhiyang He, Zhiyuan He, Zihao He, Tobias Hecking, Ming Shan Hee, Johannes Heinecke, Benjamin Heinzerling, Léo Hemamou, Ricardo Henao, James Henderson, Peter Henderson, Lisa Anne Hendricks, Iris Hendrickx, Christian Herold, Christian Heumann, Tan Hexiang, Christopher Hidey, Shohei Higashiyama, Tsutomu Hirao, Tatsuya Hiraoka, Joyce Ho, Xanh Ho, Thai Quoc Hoang, Thanh Tung Hoang, Jessica Hoffmann, Carolin Holtermann, Annie Holtz, Nils Holzenberger, Ukyo Honda, Giwon Hong, Jiajun Hong, Pengyu Hong, Seokhee Hong, Seongtae Hong, Yihuai Hong, Yu Hong, Yuan Hong, Zijin Hong, Liu Hongjun, Sara Hooker, Jacob Louis Hoover, Zachary William Hopton, Arpana Hosabetu, Sho Hoshino, Faezeh Hosseini, Sayed Mohammadreza Tayaranian Hosseini, Abe Bohan Hou, Bairu Hou, Guiyang Hou, Haowen Hou, Lu Hou, Suyang Hou, Wenjun Hou, Yifan Hou, Yingyan Hou, Yuexian Hou, Zejiang Hou, Zhaoyi Hou, Zhenyu Hou, Zhongni Hou, Eduard Hovy, Phillip Howard, Cheng-Ping Hsieh, Chun-Nan Hsu, Ting-Yao Hsu, Yu-Yin Hsu, Phu-Mon Htut, Guimin Hu, Guoping Hu, Hai Hu, Haifeng Hu, Hailin Hu, Hanxu Hu, Hongxin Hu, Jingyu Hu, Jinpeng Hu, Kaidi Hu, Lianyu Hu, Lijie Hu, Minda Hu, Minghao Hu, Nan Hu, Po Hu, Qisheng Hu, Renfen Hu, Siyuan Hu, Songbo Hu, Tiancheng Hu, Wei Hu, Wenbo Hu, Xiangkun Hu, Xiaomeng Hu, Xiaoyu Hu, Xinyu Hu, Xiping Hu, Xiyang Hu, Yebowen Hu, Yibo Hu, Yifan Hu, Yiran Hu, Yue Hu, Yushi Hu, Yuxuan Hu, Zhanghao Hu, Zhe Hu, Zhengyu Hu, Zhiyuan Hu, Zi-Yuan Hu, Ermo Hua, Nan Hua, Ting Hua, Wenyue Hua, Xian-Sheng Hua, Yilun Hua, Yuncheng Hua, Wang HuaZheng, Baoxing Huai, Baizhou Huang, Canbin Huang, Chengkai Huang, Chengsong Huang, Chengyu Huang, Chengyue Huang, Chien-yu Huang, Chu-Ren Huang, Chun-Ying Huang, Dong Huang, H. Howie Huang, Hanxian Huang, Haojing Huang, Haoyu Huang, Hen-Hsen Huang, Heng Huang, Heyan Huang, Hui Huang, James Y. Huang, Jen-tse Huang, Jerry Huang, Jimin Huang, Jimmy Huang, Jinfa Huang, Jing Huang, Jingyuan Huang, Jun Huang, Kaiyu Huang, Kaizhu Huang, Langlin Huang, Lei Huang, Liusheng Huang, Minghui Huang, Peijie Huang, Pengcheng Huang, Qingbao Huang, Qiushi Huang, Quzhe Huang, Rikui Huang, Ruixuan Huang, Shizhou Huang, Shujian Huang, Shuo Huang, Sirui Huang, Siyuan Huang, Songfang Huang, Wei Huang, Weijing Huang, Wenhao Huang, Wenxin Huang, Xiang Huang, Xiaolei Huang, Xiaowei Huang, Xu Huang, Yichong Huang, Yiming Huang, Yin Jou Huang, Yongfeng Huang, Youcheng Huang, Yu-Shiang Huang, Yuchen Huang, Yue Huang, Yufei Huang, Yuqing Huang, Yuxin Huang, Yuxuan Huang, Zanning Huang, Zengfeng Huang, Zhen Huang, Zhidian Huang, Zhiqi Huang, Zhongzhan Huang, Zijie Huang, Ziyang Huang, Yuanxiang Huangfu, Zhang Huaping, Céline Hudelot, Frederikus Hudi, Mathew John Huerta-Enochian, Stéphane Huet, Binyuan Hui, Bo Hui, Zheng Hui, Lawrence Hunter, Wenshuai Huo, Aafiya Shamshad Hussain, Kedir Yassin Hussien, Maeve Hutchinson, Nguyen Thanh Huy, Tran Ngoc Huynh, Alyssa Hwang, Dae Yon Hwang, EunJeong Hwang, Jena D. Hwang, Seonjeong Hwang, Sung Ju Hwang, Yerin Hwang, Truong Son Hy, Katharina Hämmerl, Sviatlana Höhn

Zhi Ii, Hal Daumé Iii, Taichi Iki, Nikolai Ilinykh, Irina Illina, Dmitry Ilvovsky, Alfarabi Imashev, Joseph Marvin Imperial, Muhammad Imran, Michimasa Inaba, Hirofumi Inaguma, Mert Inan, Alcides Alcoba Inciarte, Go Inoue, Koji Inoue, Ioan-Bogdan Iordache, Hasan Iqbal, David Martinez Iraola, Patrick Amadeus Irawan, Ozan Irsoy, Vatche Isahagian, Yoichi Ishibashi, Tatsuya Ishigaki, Alvi Md Ishmam, Md Asiful Islam, Mohammed Saidul Islam, Saad Obaid Ul Islam, Tunazzina Islam, Mete Ismayilzada, Paolo Italiani, Itay Itzhak, Vladimir Ivanov, Hamish Ivison, Yusuke Iwasawa, Vivek Iyer

Karthika N J, Manoj Balaji J, Tommi Jaakkola, Musashi Jacobs-Harukawa, Aditya Ajay Jadhav, Aref Jafari, Arushi Jain, Sanil Jain, Vinija Jain, Ajay Kumar Jaiswal, Terne Sasha Thorn Jakobsen,

Shruti Jalan, Dongjun Jang, Eugene Jang, Jihyoung Jang, Joonwon Jang, Sanghwan Jang, Fotis Jannidis, Peter Jansen, Arkadiusz Janz, Amit Jaspal, Sujay Kumar Jauhar, Dávid Javorský, Sai Muralidhar Jayanthi, Sébastien Jean, Nicolaas Paul Jedema, Donghyeon Jeon, Eojin Jeon, Hyewon Jeon, Yejin Jeon, Minbyul Jeong, Soyeong Jeong, Sullam Jeoung, Bin Ji, Cheng Ji, Heng Ji, Ke Ji, Lei Ji, Shengpeng Ji, Shouling Ji, Tianbo Ji, Wei Ji, Yixin Ji, Yuelyu Ji, Zhenzhou Ji, Ziwei Ji, Feiran Jia, Furong Jia, Jinghan Jia, Jinyuan Jia, Qi Jia, Weijia Jia, Xueying Jia, Yuxiang Jia, Liu Jiaheng, Ping Jian, Bin Jiang, Chao Ya Jiang, Daxin Jiang, Di Jiang, Dongfu Jiang, Fan Jiang, Fengqing Jiang, Gongyao Jiang, Guochao Jiang, Haiyun Jiang, Haoming Jiang, Huixing Jiang, Jiarong Jiang, Jinhao Jiang, Jiyue Jiang, Junfeng Jiang, Lan Jiang, Liwei Jiang, Meng Jiang, Ming Jiang, Nan Jiang, Pengcheng Jiang, Qian Jiang, Shan Jiang, Shanshan Jiang, Shengyi Jiang, Shuo-ran Jiang, Shuyang Jiang, Song Jiang, Xiaotong Jiang, Xin Jiang, Yidi Jiang, Yifan Jiang, Yong Jiang, Yu-Gang Jiang, Yukun Jiang, Yuxin Jiang, Zhengping Jiang, Zhihua Jiang, Zhongtao Jiang, Zhuoren Jiang, Zifan Jiang, Ziyang Jiang, Ziyue Jiang, Cathy Jiao, Fangkai Jiao, Jian Jiao, Wen-xiang Jiao, Xiaoqi Jiao, Yeo Wei Jie, Bowen Jin, Chuanyang Jin, Di Jin, Jiho Jin, Kyohoon Jin, Li Jin, Lisa Jin, Mingyu Jin, Renren Jin, Ruochen Jin, Ruochun Jin, Senjie Jin, Tao Jin, Xiaolong Jin, Xiaomeng Jin, Yaohui Jin, Yiqiao Jin, Yujie Jin, Zhuoran Jin, Abhishek Jindal, Disha Jindal, Ishan Jindal, Liping Jing, Liqiang Jing, Shaoling Jing, Wu Jing, Yuu Jinnai, Daejin Jo, Hwiyeol Jo, Yongrae Jo, Mayank Jobanputra, Moa Johansson, Neha Anna John, Nebojsa Jojic, Richard Adolph Aires Jonker, JiaHuei Joo, Alipio Jorge, Joemon M. Jose, Abhinav Joshi, Anant Joshi, Brihi Joshi, Kandarp Joshi, Nitish Joshi, Sachindra Joshi, Antonio Laverghetta Jr., Tianjie Ju, Xincheng Ju, Da-Cheng Juan, Swanie Juhng, Lu JunYu, Iman Jundi, Gurusha Juneja, Changsoo Jung, Dahyun Jung, Dongwon Jung, Haeji Jung, Kyomin Jung, Sangkeun Jung, Woohwan Jung, YeonJoon Jung, Li Junlin, Tom S. Juzek, Lena Ann Jäger

Mohsinul Kabir, Anubha Kabra, David Kaczér, Ram Mohan Rao Kadiyala, Marek Kadlčík, Kazuma Kadowaki, Lucie-Aimée Kaffee, Patrick Kahardipraja, Bhavya Kailkhura, Tomasz Jan Kadjanowicz, Dariusz Kajtoch, Abisek Rajakumar Kalarani, Julie Kallini, Gauri Kambhatla, Go Kamoda, Jonathan Kamp, Onno P. Kampman, Jaap Kamps, Mahammed Kamruzzaman, Deokhyung Kang, Dongjin Kang, Inho Kang, Jaewoo Kang, Jaewook Kang, Juyeon Kang, Liu Ze Kang, Min-ki Kang, Minseok Kang, Pilsung Kang, SeongKu Kang, Xiaoxi Kang, Yangyang Kang, Zhanhui Kang, Karthik reddy Kanjula, Yash Kankanampati, Evangelos Kanoulas, Yoav Kantor, Gopichand Kanumolu, Jiun-Yu Kao, Mubbasir Kapadia, Pavan Kapanipathi, Serpil Karabuklu, Antonia Karamolegkou, Mladen Karan, Shirish Karande, Amir Hossein Kargaran, Priyanka Kargupta, Rezaul Karim, Akbar Karimi, Deep Tejas Karkhanis, Bihag Karnani, Zohar Karnin, Alexey Karpov, Fakhrri Karray, Pradeep Karuturi, George Karypis, Pranav Kasela, Hisashi Kashima, Gautam Siddharth Kashyap, Zdeněk Kasner, Ali M. Kassem, Nora Kassner, Akihiko Kato, Yoshihide Kato, Shachar Katz, Manohar Kaul, Ayush Kaushal, Muhammad Dehan Al Kautsar, Parsa Kavehzadeh, Pride Kavumba, Daisuke Kawahara, Masayuki Kawarada, Hideto Kazawa, Nazmul Kazi, Przemyslaw Kazienko, Cai Ke, Chris Kedzie, Sedrick Scott Keh, Mamadou K. Keita, Katherine A. Keith, Amr Keleg, Mikaela Keller, Roman Kern, Kristian Kersting, Natthawut Kertkeidkachorn, Vaishakh Keshava, Baber Khalid, Muhammad Khalid, Muhammad Khalifa, Zena Al Khalili, Mohammed Khalilia, Fahad Shahbaz Khan, Latifur Khan, Mohammad Aflah Khan, Aditi Khandelwal, Shima Khanehzar, Simran Khanuja, Subhojyoti Khastagir, Faiza Khan Khattak, Yassine El Kheir, Vivek Khetan, Hichem Ammar Khodja, Urja Khurana, Dayeon Ki, Douwe Kiela, Johannes Kiesel, Zlata Kikteva, Bugeun Kim, Byeonggeun Kim, Byoungjip Kim, Chae Won Kim, Chaeun Kim, Changhun Kim, Dahyun Kim, Donghyun Kim, Doyoung Kim, Eunji Kim, Gangwoo Kim, Geewook Kim, Gene Louis Kim, Gunhee Kim, Gyuwan Kim, Hannah Kim, Harksoo Kim, Hazel H. Kim, Heeseung Kim, Hongjin Kim, Hwichan Kim, Hyuhng Joon Kim, Jaehyung Kim, Jaein Kim, Jaeyoung Kim, Jeonghwan Kim, Ji-Hoon Kim, Jieyong Kim, Jihie Kim, Jiho Kim, Jihyuk Kim, Jinsung Kim, Jiseon Kim, Jongyoon Kim, Juho Kim, Jung-jae Kim, Junho Kim, Junmo Kim, Junyeob Kim, Junyeong Kim, Kang-Min Kim, Kangil Kim, Michelle Kim, Minbeom Kim, Minseok Kim,

Minsoo Kim, Minsu Kim, Misuk Kim, Namyoung Kim, Seungone Kim, Sun Kim, Sungchul Kim, Sungwoong Kim, Taehee Kim, Taehwan Kim, Taesup Kim, Yekyung Kim, Yongil Kim, Youna Kim, YoungBin Kim, Youngkyoung Kim, Youngwoo Kim, Yujin Kim, Yumin Kim, Yunsoo Kim, Zae Myung Kim, Vadim Kimmelman, Brendan King, Tracy Holloway King, Yuval Kirstain, Frederic Kirstein, Varsha Kishore, Diego Klabjan, Jens Kleesiek, Ayal Klein, Dorde Klisura, Alois Knoll, Jongwoo Ko, Sang-Ki Ko, Hideo Kobayashi, Ichiro Kobayashi, Elena Kochkina, Jan Kocon, Prashant Kodali, Joachim Koehler, Hyukhun Koh, Ryosuke Kohita, Guneet Singh Kohli, Takeshi Kojima, Stanley Kok, Olga Kolesnikova, Prateek Kolhar, Alexander Koller, Boshko Koloski, Rik Koncel-Kedzior, Grzegorz Kondrak, Cunliang Kong, Fang Kong, Fanshuang Kong, Jun Kong, Keyi Kong, Lingxing Kong, Vasily Konovalov, Myoung-Wan Koo, Seonmin Koo, Bevan Koopman, Yeganeh Kordi, Mandy Barrett Korpusik, Katerina Korre, Roman Koshkin, Satoshi Kosugi, Katsunori Kotani, Neema Kotonya, Ziyi Kou, Vasiliki Kougia, George Kour, Punit Singh Koura, Ross Koval, Adriana Kovashka, Ivan Koychev, Oluwasanmi O Koyejo, Elisa Kreiss, Matt Kretchmar, Ranjay Krishna, Satyapriya Krishna, Lun-Wei Ku, Max Ku, Jiayi Kuang, Kun Kuang, Marek Kubis, Keito Kudo, Sachit Kuhar, Marco Kuhlmann, Ilia Kulikov, Adithya Kulkarni, Atharva Kulkarni, Mayank Kulkarni, Sayali Kulkarni, Artur Kulmizev, Saurabh Kulshreshtha, Anoop Kumar, Bhargava Kumar, Karun Kumar, Prince Kumar, Ravi Kumar, Rishabh Kumar, Sandeep Kumar, Shachi Hullumane Kumar, Shivani Kumar, Sonal Kumar, Srijan Kumar, Sujay S Kumar, Varun Kumar, Vineet Kumar, Vivek Kumar, Yaman Kumar, Ponnurangam Kumaraguru, Gitanjali Kumari, Tuhin Kundu, Po-Nien Kung, Olli Kuparinen, Kemal Kurniawan, Sadao Kurohashi, Venkata Madhu Sravanth Kurumaddali, Wojciech Kusa, Saar Kuzi, Gleb Kuzmin, Andrey Kuznetsov, Ilia Kuznetsov, Kristian Kuznetsov, Dong-Hyun Kwak, Haewoon Kwak, Nojun Kwak, Wonbin Kweon, James Kwok, Kenneth Kwok, Deuksin Kwon, Jingun Kwon, Sunjae Kwon, Tanja Käser, Abdullatif Köksal

Jorma Tapio Laaksonen, Gorka Labaka, Cyril Labbe, Yanis Labrak, Faisal Ladhak, Avishek Lahiri, Allison Claire Lahnala, Catherine Lai, Hanjiang Lai, Huiyuan Lai, Kenneth Lai, Peichao Lai, Viet Dac Lai, Wen Lai, Yi-An Lai, Yilong Lai, Yuhang Lai, Kushnareva Laida, Nurkhan Laiyk, Vasudev Lal, Yash Kumar Lal, John P. Lalor, Albert Y.S. Lam, Tsz Kin Lam, Hemank Lamba, Nathan Lambert, Andrew Lan, Man Lan, Tian Lan, Yifan Lan, Zhenzhong Lan, Zhibin Lan, Hao Lang, Qiwei Lang, Karin De Langis, David Langlois, Curtis Langlotz, Mateusz Lango, Gabriella Lapesa, Christine Largeron, Daniil Larionov, Sahinur Rahman Laskar, Gregory Kang Ruey Lau, Jey Han Lau, Alon Lavie, Neal Gregory Lawton, Dung Duy Le, Duy Le, Hang Le, Nghia T. Le, Phong Le, Trung Le, Joshua Ong Jun Leang, Gwénolé Lecorvé, Bruce W Lee, Byung-Jun Lee, Celine Lee, Changki Lee, Changmin Lee, Chen-Yu Lee, Cheryl Lee, Chia-Jung Lee, Daniel Lee, Dohyeon Lee, Dohyun Lee, Dongha Lee, Dongryeol Lee, Dongwon Lee, Fei-Tzin Lee, Gyubok Lee, Hojae Lee, Honglak Lee, Hwiwon Lee, Hyunji Lee, Jae Hee Lee, Jaejin Lee, Jaeseong Lee, Jaewook Lee, Jaeyoung Lee, Jimin Lee, Jinu Lee, Jongwuk Lee, Jungseob Lee, Junho Lee, Junhyeok Lee, Kang-il Lee, Kyogu Lee, Kyumin Lee, Mingyu Lee, Minhwa Lee, Noah Lee, Saehyung Lee, Sang-goo Lee, Sanwoo Lee, Seungeon Lee, Seungyeon Lee, Sunkyung Lee, Wonkee Lee, Yoonjoo Lee, Young-Jun Lee, Younghun Lee, Youngwon Lee, Yukyung Lee, Els Lefever, Jens Lehmann, Deren Lei, Fangyu Lei, Shun Lei, Xuanyu Lei, Yibin Lei, Zhenyu Lei, Markus Leipold, Christoph Leiter, Gaël Lejeune, Leon Leng, Chak Tou Leong, Colin Leong, Bruno Lepri, Pietro Lesci, Gina-Anne Levow, Sharon Levy, Martha Lewis, Patrick Lewis, Ang Li, Bangzheng Li, Baochun Li, Baohang Li, Bin Li, Bing Li, Binyang Li, Bo Li, Bobo Li, Bocheng Li, Boxuan Li, ChangHao Li, Changmao Li, Changqun Li, Changye Li, Chaozhao Li, Chengkai Li, Chengming Li, Chengpeng Li, Chengxi Li, Chengzhengxu Li, Chengzu Li, Chenliang Li, Chong Li, Chunyang Li, Chuyuan Li, Cuiping Li, Daifeng Li, Dawei Li, Dexun Li, Deyu Li, Dingcheng Li, Diya Li, Dong Li, Dongsheng Li, Dongyang Li, Dongyuan Li, Feng-Lin Li, Gang Li, Ge Yi Li, Haizhou Li, Hanbo Li, Hanqi Li, Hao Li, Haochen Li, Haohang Li, Haoran Li, Haoxin Li, Haoyuan Li, He Li, Huayang Li, Hui Li, Huihan Li, Huiying Li, Jia Li, Jiaang Li, Jiachen Li, Jian Li, Jiang Li,



Jianjun Li, Jianling Li, Jianxin Li, Jiaoda Li, Jiaqi Li, Jiaxi Li, Jiazhao Li, Jiazheng Li, Jierui Li, Jijie Li, Jinyang Li, Jinyuan Li, Jixing Li, Juan Li, Juanhui Li, Juanzi Li, Judith Yue Li, Juncheng Li, Jundong Li, Juntao Li, Junzhuo Li, Kaiyuan Li, Kan Li, Ke Li, Kun Li, Lei Li, Li Li, Liang Li, Liangyue Li, Liangze Li, Lihong Li, Lijun Li, Lin Li, Linjing Li, Linjun Li, Linlin Li, Lishuang Li, Liyao Li, Lujun Li, Mengze Li, Miao Li, Miaomiao Li, Min Li, Ming Li, Mingchen Li, Mingming Li, Mingzhe Li, Mukai Li, Muzhi Li, Nian Li, Panfeng Li, Peiguang Li, Peng Li, Ping Li, Qi Li, Qian Li, Qing Li, Qintong Li, Qiuchi Li, Ran Li, Raymond Li, Renhao Li, Rongjun Li, Rongsheng Li, Rui Li, Ruifan Li, Ruiqi Li, Ruixuan Li, Rumei Li, Ruochen Li, Ruosen Li, Ruoyu Li, Ryan Li, Senyu Li, Sha Li, Shang-Wen Li, Shaobo Li, Shasha Li, Sheng Li, Shengjie Li, Shichen Li, Shimin Li, Shiyang Li, Shuangyin Li, Shuo Li, Shuyue Stella Li, Sihang Li, Siheng Li, Siyan Li, Sujian Li, Tianhao Li, Tianjian Li, Tianlin Li, Tianrui Li, Tianshi Li, Victor O.K. Li, Wei Li, Weikang Li, Weiping Li, Wen-Ding Li, Wenda Li, Wenhao Li, Wenyan Li, Xia Li, Xian Li, Xiang Li, Xiang Lisa Li, Xiangci Li, Xianming Li, Xiaochuan Li, Xiaonan Li, Xiaopeng Li, Xin Li, Xingxuan Li, Xinjin Li, Xinlin Li, Xintong Li, Xinze Li, Xuchen Li, Xurui Li, Yang Li, Yangning Li, Yanyang Li, Yi Li, Yifei Li, Ying Li, Yinghao Li, Yingya Li, Yitong Li, Yiwei Li, Yixia Li, Yiyang Li, Yizhi Li, Yong Li, Yongqi Li, Yu Li, Yuan Li, Yucheng Li, Yudong Li, Yuepei Li, Yun Li, Yunpeng Li, Yunshui Li, Yuntao Li, Yunxin Li, Yunxuan Li, Yuqing Li, Zang Li, Zechen Li, Zehan Li, Zekun Li, Zexin Li, Zexuan Li, Zhaocong Li, Zhaonan Li, Zhaowei Li, Zhaoyi Li, Zhen Li, Zheng Li, Zhenwen Li, Zhixu Li, Zhoujun Li, Zhuochun Li, Zhuohang Li, Zhuohao Li, Zhuoqun Li, Zhuoran Li, Zhuowan Li, Zhuoyan Li, Zichao Li, Zihan Li, Zihao Li, Zizhong Li, Zongxi Li, Zongxia Li, Zongyao Li, Zuchao Li, Defu Lian, Jianxun Lian, Chen Liang, Davis Liang, Jianxin Liang, Jiaqing Liang, Jinggui Liang, Jiye Liang, Juhao Liang, Ke Liang, Lei Liang, Shining Liang, Shuo Liang, Sichu Liang, Tian Liang, Tingting Liang, Xiaobo Liang, Xiaozhuan Liang, Xiwen Liang, Yingyu Liang, Yitao Liang, Yunlong Liang, Yuxuan Liang, Zhenwen Liang, Baohao Liao, Hao Liao, Jian Liao, Lei Liao, Minpeng Liao, Ruotong Liao, Xixian Liao, Yong Liao, Yusheng Liao, Zihan Liao, Luca Di Liello, Erin MacMurray van Liemt, Wencke Liermann, Veronica Liesaputra, Jasy Suet Yan Liew, Gilbert Lim, Jungwoo Lim, KyungTae Lim, Ser-Nam Lim, Woohyung Lim, Woosang Lim, Zheng Wei Lim, Peerat Limkonchotiawat, Antonie Lin, Binghuai Lin, Bingqian Lin, Chen Lin, Chenghua Lin, Dahua Lin, Fangquan Lin, Gang Lin, Haokun Lin, Haonan Lin, Hongzhan Lin, Huawei Lin, I-Fan Lin, Inna Wanyin Lin, Jiayu Lin, Jieru Lin, Jiuheng Lin, Jun Lin, Kevin Lin, Liang Lin, Lu Lin, Mingbao Lin, Minhua Lin, Pin-Jie Lin, Qian Lin, Qika Lin, Qinghong Lin, Qingwei Lin, Run Lin, Sheng-Chieh Lin, Shom Lin, Shuhang Lin, Tony Lin, Tzu-Quan Lin, Weiran Lin, Weizhe Lin, Xin Alex Lin, Xuemin Lin, Xueyuan Lin, Yen-Ting Lin, Yi-Cheng Lin, Yijie Lin, Ying Lin, Youyuan Lin, Yueqian Lin, Zeqi Lin, Zhe Lin, Zhen Lin, Zheng Lin, Zhenxi Lin, Zhouhan Lin, Zi Lin, Zihao Lin, Zijia Lin, Chen Ling, Zhen-Hua Ling, Zijian Ling, Pietro Lio, Gili Lior, Zachary Lipton, Huang Lisheng, Aiwei Liu, Anqi Liu, Ben Liu, Bin Liu, Bo Liu, Bobobo Liu, Chang Liu, Chen Cecilia Liu, Chengyuan Liu, Chuanyi Liu, Chunhua Liu, Cong Liu, Daizong Liu, Danni Liu, Dayiheng Liu, Deyuan Liu, Dongfang Liu, Dongrui Liu, Fangchao Liu, Fei Liu, Fuxiao Liu, Gongshen Liu, Guangyi Liu, Guojun Liu, Han Liu, Hang Liu, Hanmeng Liu, Hao Liu, Heyang Liu, Hongfu Liu, Hui Liu, Huijun Liu, Jason Klein Liu, Ji Liu, Jia Liu, Jiacheng Liu, Jiachi Liu, Jiale Liu, Jiamou Liu, Jianfeng Liu, Jiaqing Liu, Jiarui Liu, Jiateng Liu, Jiawei Liu, Jiaxiang Liu, Jie Liu, Jie-jyun Liu, Jin Liu, Jing Liu, Jinghui Liu, Jingping Liu, Jinxin Liu, Jiongnan Liu, Juhua Liu, Jun Liu, Junnan Liu, Kang Liu, Kuanghong Liu, Kunpeng Liu, Lihui Liu, Lin Liu, Meng Liu, Ming Liu, Naiming Liu, Nayu Liu, Ning Liu, Ninghao Liu, Peiyu Liu, Pengyuan Liu, Qiang Liu, Qiao Liu, Qijiong Liu, Qin Liu, Qingbin Liu, Qinqing Liu, Quan Liu, Rui Liu, Ruibo Liu, Ruiheng Liu, Ruikang Liu, Ruitong Liu, Ruiyang Liu, Shangqing Liu, Sheng Liu, Shenghua Liu, Shengjie Liu, Shengping Liu, Shicheng Liu, Shilei Liu, Shiwei Liu, Shuheng Liu, Shujie Liu, Shuliang Liu, Shuqi Liu, Sijia Liu, Siyang Liu, Siyi Liu, Tianci Liu, Tianqiao Liu, Tianyang Liu, Ting Liu, Tong Liu, Wanlong Liu, Wei Liu, Weifeng Liu, Weisi Liu, Wenhe Liu, Wenrui Liu, Xiang Liu, Xianggen Liu, Xianglong Liu, Xiaoming Liu, Xiaoqian Liu, Xiaoran Liu, Xiaozhong Liu, Xin Liu, Xinyu Liu, Xinyue Liu, Xixian Liu, Xubo Liu, Xue Liu, Xuebing Liu,

Xukai Liu, Xuye Liu, Yahui Liu, Yan Liu, Yanchen Liu, Yanchi Liu, Yang Liu, Yang Aron Liu, Yanming Liu, Yantao Liu, Yao Liu, Ye Liu, Yi Liu, Yicun Liu, Yihong Liu, Yilun Liu, Yiming Liu, Yinhong Liu, Yong Liu, YongKang Liu, Yongbin Liu, Yonghao Liu, Yongmei Liu, Yu Liu, Yuan Liu, Yuanxing Liu, Yue Liu, Yufang Liu, Yuhan Liu, Yunxin Liu, Yurong Liu, Zeming Liu, Zhengyuan Liu, Zhengzhong Liu, Zhenhua Liu, Zhexiong Liu, Zheyuan Liu, Zhi Liu, Zhili Liu, Zhining Liu, Zhiwei Liu, Zhiyue Liu, Zihan Liu, Zihang Liu, Zijie Liu, Zijing Liu, Zijun Liu, Zikang Liu, Ziqi Liu, Zirui Liu, Zitao Liu, Ziyi Liu, Soda Marem Lo, Tien-Hong Lo, Tyler Loakman, Colin Lockard, Mengsay Loem, Le Huu Loi, Rahul Lokesh, Maria Lomeli, Do Xuan Long, Jikai Long, Qingqing Long, Quanyu Long, Teng Long, Wanqiu Long, Xinwei Long, Yunfei Long, Michela Lorandi, Ehsan Lotfi, Cedric Lothritz, Chao Lou, Qian Lou, Natalia Loukachevitch, Bo-Ru Lu, Chaochao Lu, Heng-yang Lu, Jianqiao Lu, Jiasen Lu, Jinghui Lu, Jinliang Lu, Junru Lu, Keming Lu, Qingyu Lu, Rongxing Lu, TaiMing Lu, Tianyu Lu, Wenpeng Lu, Xiaoxin Lu, Xin Lu, Xingyu Lu, Xinjiang Lu, Xinyuan Lu, Xuesong Lu, Yao Lu, Yunan Lu, Yuxuan Lu, Yuyin Lu, Zexin Lu, Jian Luan, Evan Lucas, Jason S Lucas, Bernd Ludwig, Agnes Luhtaru, Stephanie M. Lukin, Chen Luo, Chu Fei Luo, Fuwen Luo, Haoran Luo, Hongyin Luo, Jiebo Luo, Jinqi Luo, Junyu Luo, Kangyang Luo, Lannan Luo, Ling Luo, Linhao Luo, Man Luo, Minnan Luo, Ruilin Luo, Sichun Luo, Siwen Luo, Tianze Luo, Weidi Luo, Xiang Luo, Xianzhen Luo, Yang Luo, Yao Luo, Yingfeng Luo, Yiran Luo, Yuanchang Luo, Yuyu Luo, Zheheng Luo, Zhuang Luo, Zhunchen Luo, Ziming Luo, Thang Luong, Marlene Lutz, Lin Luyang, Bo Lv, Changze Lv, Fengmao Lv, Kai Lv, Xingtai Lv, Zheqi Lv, Kevin Lybarger, Chris Lydia, Bohan Lyu, Chengqi Lyu, Chenyang Lyu, Chunchuan Lyu, Lingjuan Lyu, Shengfei Lyu, Xinglin Lyu, Xinxi Lyu, Yougang Lyu, Zongwei Lyu, Alexander Lyzhov, Zhiyu li

Bo Ma, Bolei Ma, Chen Ma, Congbo Ma, Fukun Ma, Guangyuan Ma, Jiawei Ma, Jie Ma, Jun-Yu Ma, Lei Ma, Longxuan Ma, Pingchuan Ma, Qianli Ma, Rao Ma, Shiqing Ma, Shuo Ma, Simin Ma, Tengfei Ma, Weicheng Ma, Weitao Ma, Weizhi Ma, Xinbei Ma, Xinge Ma, Xinyin Ma, Xueguang Ma, Xutai Ma, Yanan Ma, Yichuan Ma, Yunpu Ma, Yuqing Ma, Zexiong Ma, Zhengrui Ma, Zihan Ma, Ziyang Ma, Iffat Maab, Kate MacFarlane, Sashank Macha, Lieve Macken, Avinash Madasu, Mounica Maddela, Rahul Madhavan, Sathwik Tejaswi Madhusudhan, Manas Madine, Tharindu Madusanka, Seiji Maekawa, Joao Magalhaes, Samar Mohamed Magdy, Walid Magdy, Bernardo Magnini, Måns Magnusson, Ahmed Magooda, Shweti Mahajan, Kishan Maharaj, Ridwan Mahbub, Rahmad Mahendra, A V S D S Mahesh, Ayush Maheshwari, Rishabh Maheshwary, Florian Mai, Sijie Mai, Mieradilijiang Maimaiti, Navonil Majumder, Brian Mak, Márton Makrai, Emanuele La Malfa, Mahan Malihi, Manuj Malik, Vijit Malik, Bradley A. Malin, Kolya Malkin, Christopher Malon, Yash Malviya, Zhibo Man, Sunny Manchanda, Amogh Mannekote, Christopher D. Manning, Dinesh Manocha, Ioana Manolescu, Irene L Manotas, Ramesh Manuvinakurike, Muhammad Arslan Manzoor, Cunli Mao, Shaoguang Mao, Xiaoxi Mao, Yuzhen Mao, Zhiming Mao, Zihao Mao, Ana Marasovic, Guillermo Marco, Marion Di Marco, Alda Mari, Vukosi N. Marivate, Ilia Markov, Magdalena Markowska, Marc Marone, Lluís Marquez, Mounika Marreddy, Jason Mars, Federico Martelli, Anna Martin, James Martin, Maria-teresa Martin-valdivia, Matej Martinc, Marianna J. Martindale, Andre Martins, Bruno Martins, Antonio David Ponce Martínez, Mihai Masala, Reem Masoud, Shahed Masoudian, Ahmed Masry, Sarah Masud, Marco Matasconi, Ved Mathai, Brodie Mather, Leena Mathur, Puneet Mathur, Shivam Mathur, Yossi Matias, Glenn Matlin, David Martins de Matos, Yuji Matsumoto, Yutaka Matsuo, Yevgen Matuskevych, Evgeny Matusov, Kaushal Kumar Maurya, Costas Mavromatis, Chandler May, Philipp Mayr, Afshaan Mazagonwalla, Sabrina McCallum, John Philip McCrae, Kate McCurdy, Daniel McDuff, Euan McGill, Barbara McGillivray, Nick McKenna, Paul McNamee, Sourav Medya, Michiel van der Meer, Karine Megerdooian, Alexander Mehler, Shuhaib Mehri, Ambuj Mehrish, Atharva Mehta, Kartik Mehta, Maitrey Mehta, Rahul Mehta, Sanket Vaibhav Mehta, Sushant Mehta, Jingbiao Mei, Kai Mei, Lingrui Mei, Sen Mei, Stephen Meisenbacher, Abdellah El Mekki, Jesus Lovón Melgarejo, Gerard Melo, Ethan Adrian Mendes, John Mendonca, Chuan Meng, Fanchao

Meng, Helen M. Meng, Shiao Meng, Yen Meng, Yu Meng, Yuanliang Meng, Ziqiao Meng, Stefano Menini, Robert Mercer, Fabio Mercorio, Mike A Merrill, William Merrill, Rin Metcalf, Eleni Metheniti, Guillaume Metzler, Selina Meyer, Fei Mi, Md Messal Monem Miah, Yibo Miao, Yisong Miao, Yuchun Miao, Kurt Micallef, Andrianos Michail, Zulfat Miftahutdinov, Naiara Perez Miguel, Matvey Mikhalechuk, Jeremiah Milbauer, Filip Miletic, Evangelos Milios, Isaac Miller, Timothy A. Miller, David Mimno, Dehai Min, Do June Min, Qingkai Min, Pasquale Minervini, Koji Mineshima, Chenlin Ming, Benjamin Minixhofer, Einat Minkov, Hideya Mino, Lester James Validad Miranda, Santiago Miret, Seyedabolghasem Mirroshandel, Mehrnoosh Mirtaheri, Katsiaryna Mirylenka, Paramita Mirza, Anand Mishra, Kshitij Mishra, Prakamya Mishra, Pruthwik Mishra, Swaroop Mishra, Amita Misra, Cassie S. Mitchell, Avijit Mitra, Chancharik Mitra, Pabitra Mitra, Subrata Mitra, Ashish Mittal, Yusuke Miyao, Taro Miyazaki, Tong Mo, Wenjie Mo, Ali Modarressi, Arkadiusz Modzelewski, Lucas Moeller, Abdelrahman Mohamed, Amr Mohamed, Youssef Sherif Mansour Mohamed, Nabeel Mohammed, Biswesh Mohapatra, Fedor Moiseev, Jisoo Mok, Francesco Maria Molfese, Nicholas Monath, Ishani Mondal, Mainack Mondal, Philipp Mondorf, Joel Ruben Antony Moniz, Syed Mostofa Monsur, Arturo Montejó-Ráez, Sebastien Montella, Kyle Montgomery, Andrés Montoyo, Christof Monz, Hyeonseok Moon, Sangwhan Moon, Seungwhan Moon, Ray Mooney, Guy Mor, Raha Moraffah, Victor Morand, Tatsunori Mori, Emmanuel Morin, Maria Moritz, Robert Moro, Luca Moroni, Jacob Daniel Morrison, Viktor Moskvoretskii, Yasmin Moslem, Josiane Mothe, Xinyi Mou, Yutao Mou, Noura Al Moubayed, Basel Ahmad Mousi, Li Moxin, Jamshid Mozafari, Yongyu Mu, Hamdy Mubarak, David Mueller, Niklas Muennighoff, Animesh Mukherjee, Anjishnu Mukherjee, Arjun Mukherjee, Sagnik Mukherjee, Tanmoy Mukherjee, Matthew Mulholland, Sahal Shaji Mullappilly, Carlos Mullov, Asim Munawar, Rajesh Kumar Mundotiya, Vishvak Saivenkat Murahari, Koji Murakami, Soichiro Murakami, Kashyap Coimbatore Murali, Masayasu Muraoka, Yugo Murawaki, Vanessa Murdock, Gabriel Murray, Rudra Murthy, Shikhar Murty, John Murzaku, Elio Musacchio, Vinod Muthusamy, Arianna Muti, Alberto Muñoz-Ortiz, Skatje Myers, Bernd Möbius, Sebastian Möller, Max Mühlhäuser

CheolWon Na, Hongbin Na, Seung-Hoon Na, Nona Naderi, Mohamed Nadif, Aliakbar Nafar, Shubhadip Nag, Manish Nagireddy, Arsha Nagrani, Md Mahadi Hasan Nahid, Atharva Naik, Inderjeet Jayakumar Nair, Shoichi Naito, Saeed Najafi, Mahyar Najibi, Tetsuji Nakagawa, Yukiko Nakano, Itay Nakash, Hootan Nakhost, Eric Nalisnick, Anugunj Naman, Mahdi Namazifar, Guoshun Nan, Yiyang Nan, Srinivasan Nandakumar, Arghodeep Nandi, Tarek Naous, Diane Napolitano, Anjali Narayan-Chen, Daniele Nardi, Yaswanth Narsupalli, Neha Narwal, Sudip Kumar Naskar, Anandhavelu Natarajan, Abhijnan Nath, Swaroop Nath, Costanza Navarretta, Nihal V. Nayak, Shravan Nayak, Tapas Nayak, Renato Negrinho, James O' Neill, Willie Neiswanger, Wolfgang Nejd, Anil Kumar Nelakanti, Poli Nemkova, Clement Neo, Vera Neplenbroek, Mariana Neves, Jennifer Neville, Raymond T. Ng, See-Kiong Ng, York Hay Ng, Youyang Ng, Henry Nghiem, Chong Wah Ngo, Nghia Trung Ngo, Axel Ngonga, Anh Totti Nguyen, Bang Nguyen, Bao Nguyen, Chien Van Nguyen, Duc-Vu Nguyen, Duy Nguyen, Huy Nguyen, Khanh Xuan Nguyen, Kiet Van Nguyen, Long Hb Nguyen, Luan Thanh Nguyen, Ngan Luu-Thuy Nguyen, Phi Xuan Nguyen, Tai Nguyen, Thong Nguyen, Tin Trung Nguyen, Tung Nguyen, Vincent Nguyen, Hoang-Quoc Nguyen-Son, John Nham, Jianmo Ni, Jingwei Ni, Xuanfan Ni, Yuansheng Ni, Keanu Nichols, Lionel Nicolas, Vlad Niculae, Dong Nie, Ercong Nie, Jian-Yun Nie, Liqiang Nie, Ping Nie, Yuxiang Nie, Zhijie Nie, Elizabeth Kaye Nielsen, Shubham Kumar Nigam, Hellina Hailu Nigatu, Malvina Nikandrou, Anna Nikiforovskaya, Irina Nikishina, Christina Niklaus, Giannis Nikolentzos, Vassilina Nikoulina, Iftitahu Nimah, Jinzhong Ning, Ruoxi Ning, Yue Ning, Kyosuke Nishida, Noriki Nishida, Satoshi Nishida, Yuto Nishida, Sergiu Nisioi, Malvina Nissim, Vikram Nitin, Hao Niu, Jianwei Niu, Jingcheng Niu, Xing Niu, Xinyuan Niu, Xueyan Niu, Yingjie Niu, Zhendong Niu, Joakim Nivre, Ayana Niwa, Bill Noble, Rodrigo Frassetto Nogueira, Tadashi Nomoto, Rik van Noord, Enrique Noriega-Atala, Kai North, Jekaterina Novikova, Robert D Nowak, Sarana Nutanong, Claire Nédellec, Tianhao niu

Tim Oates, Daisuke Oba, Nicolás Benjamín Ocampo, Daniela Occhipinti, Brendan Oconnor, Yusuke Oda, Stephan Oepen, Michael Ogezi, Odunayo Ogundepo, Dongsuk Oh, Hongseok Oh, Jaehoon Oh, Tae-hyun Oh, Yerim Oh, Jungseul Ok, Tatsushi Oka, Shu Okabe, Tsuyoshi Okita, Akintunde Oladipo, Tobi Olatunji, Fredrik Olsson, Kai Tzu-iunn Ong, Yew Soon Ong, Yasumasa Onoe, Nelleke Oostdijk, Subba Reddy Oota, Gustaw Opielka, Juri Opitz, Shereen Oraby, Jorge Orbay, Riccardo Orlando, Matthias Orlikowski, John D Osborne, Ivan Oseledets, Malte Ostendorff, Yulia Otmakhova, Jiefu Ou, Weihua Ou, Yixin Ou, Zhijian Ou, Hiroki Ouchi, Siqi Ouyang, Siru Ouyang, Wanli Ouyang, Yang Ouyang, Zhongyu Ouyang, Arnold Overwijk, Momose Oyama, Masafumi Oyamada, Savas Ozkan, Şaziye Betül Özateş, Arzucan Özgür

Aswini Kumar Padhi, Trilok Padhi, Tejomay Kishor Padole, Tim Paek, Aline Paes, Artidoro Pagnoni, Khushbu Pahwa, Vasile Pais, Christopher J Pal, Hamid Palangi, Chester Palen-Michel, Girish Keshav Palshikar, Changzai Pan, Deng Pan, Junting Pan, Leyi Pan, Rui Pan, Ruihao Pan, Shirui Pan, Xiang Pan, Xiao Pan, Xiaoman Pan, Artemis Panagopoulou, Ruchi Panchanadikar, Alexander Panchenko, Pranoy Panda, Nilay Pande, Punya Syon Pandey, Saurabh Kumar Pandey, Lalchand Pandia, Shrey Pandit, Bo Pang, Zi Haur Pang, Aleksandr Panov, Maxim Panov, George Pantazopoulos, Haris Papageorgiou, Evangelos E. Papalexakis, Paolo Papotti, Duccio Pappadopulo, Nikolaos Pappas, Georgios Paraskevopoulos, Amit Parekh, Tanmay Parekh, Angelina Parfenova, Priyank Parikh, Rahil Parikh, Christopher Parisien, Beomchan Park, ChaeHun Park, Chanyoung Park, Cheonbok Park, Cheoneum Park, Choonghyun Park, Eunhyeok Park, Gyeong-Moon Park, Gyutae Park, Hae Won Park, Hun Park, Hyuntae Park, Jae-Woo Park, Jungyeul Park, Seunghyun Park, Shinwoo Park, Sungjin Park, Sungwon Park, Woomyoung Park, Mihir Parmar, Denis Parra, Pranjali Ajay Parse, Gabriella Pasi, Nivranshu Pasricha, Rebecca Passonneau, Clemente Pasti, Valeria Pastorino, Hitesh Laxmichand Patel, Maitreya Patel, Barun Patra, Ioannis Patras, Jasabanta Patro, Priyaranjan Pattanayak, Parth Patwa, Manasi Patwardhan, Mostofa Patwary, Indraneil Arun Paul, Shounak Paul, Matthias Paulik, Adam Pauls, Vera Pavlova, Sachin Pawar, Ali Payani, Branislav Pecher, Jiahuan Pei, Georgios Peikos, Guo Peiming, Lotem Peled-Cohen, Kellin Pelrine, Abhirama Subramanyam Penamakuri, Baolin Peng, Dezhi Peng, Dezhong Peng, Jingyu Peng, Keqin Peng, Letian Peng, Miao Peng, Min Peng, Qiwei Peng, Qiyao Peng, Shuyuan Peng, Tianhao Peng, Tianshuo Peng, Wen-Chih Peng, Xi Peng, Yifan Peng, Gerald Penn, Nicolò Penzo, Valerio Pepe, Joseph Joshua Peper, Jesus Javier Calleja Perez, Carla Perez-Almendros, Laura Perez-Beltrachini, Francesco Periti, Stefano Perrella, Denis Peskov, Ben Peters, Dominic Petrak, Maxime Peyrard, Pouya Pezeshkpour, Chau Thi Minh Pham, Minh Vu Pham, Quang Hieu Pham, Van-Thuy Phi, Fred Philippy, Anirudh Phukan, Renjie Pi, Cesare Spinoso-Di Piano, Pablo Piantanida, Andrea Piergentili, Sara Pieri, Reshmi Gopalakrishna Pillai, Wei Ping, Claudio Pinhanez, Juan Pino, Andrew Piper, Stelios Piperidis, Robinson Piramuthu, Dhivya Piraviperumal, Bhawna Piryani, Dina Pisarevskaya, Hugo Pitorro, Priya Pitre, Lidia Pivovarova, Paul Piwek, Barbara Plank, Laura Plaza, Moritz Plenz, Sergey Pletenev, Esther Ploeger, Max Ploner, Soham Poddar, Christian Poelitz, Wessel Poelman, Thierry Poibeau, Rhitabrat Pokharel, Sandeep Polisetty, Senja Pollak, Peter Polák, José Pombal, Elinor Poole-Dayana, Lucian Popa, Adrian Popescu, Maxim Popov, Maja Popovic, Nicholas Popovic, Beatrice Portelli, Matt Post, Martin Potthast, Amrit Poudel, Mohammadreza Pourreza, Yi Zhang Pp, Akshara Prabhakar, Ritu Prajapati, Nirmalendu Prakash, Aniket Pramanick, A Pranav, Archiki Prasad, Shivika Prasanna, Gabriele Prato, Ian Pratt-Hartmann, Frederic Precioso, Judita Preiss, Damith Premasiri, Priyanshu Priya, Lorenzo Proietti, Prokopis Prokopiadis, Irina Proskurina, Emily Prud'hommeaux, Xiao Pu, Giovanni Puccetti, Giulia Pucci, Ratish Puduppully, Haritz Puerto, Saurabh Pujar, Rajkumar Pujari, Siddhartha Pullakhandam, Sukannya Purkayastha, Alberto Purpura, Matthew Purver, Ayu Purwarianti, Rifki Afina Putri, Prakash Reddy Putta, Juan Antonio Pérez-Ortiz, Siddhesh pawar

Eliver Q, Abdelrahim Qaddoumi, Biqing Qi, Fanchao Qi, Ji Qi, Jianzhong Qi, Jirui Qi, Men-

gnan Qi, Peng Qi, Qi Qi, Siya Qi, Tao Qi, Weihong Qi, Weijian Qi, Weizhen Qi, Yunjia Qi, Dong Qian, Jingwen Qian, Junlang Qian, Kun Qian, Shenbin Qian, Weining Qian, Xueming Qian, Yao Qian, Yining Qian, Yusu Qian, Zhong Qian, Jipeng Qiang, Kangyu Qiao, Shanbao Qiao, Shuofei Qiao, Zile Qiao, Bing Qin, Bowen Qin, Chengwei Qin, Guanghui Qin, Jing Qin, Jinghui Qin, Tian Qin, Yang Qin, Yao Qin, Yifang Qin, Yong Qin, Zhenyue Qin, Peijun Qing, Huachuan Qiu, Kaitao Qiu, Linlu Qiu, Mengyang Qiu, Shuwen Qiu, Xihe Qiu, Xin Ying Qiu, Yifu Qiu, Zexuan Qiu, Zhaopeng Qiu, Zihan Qiu, Zitai Qiu, Muhammad Reza Qorib, Dan Qu, Jianfeng Qu, Xiaoye Qu, Xingwei Qu, Yincen Qu, Tho Quan, Xin Quan, Hieu Ta Quang, Ariadna Quattoni, Solen Quiniou, Chris Quirk, Rizwan Qureshi

Gowtham R, Baraniuk R., Reihaneh Rabbany, Hamid R. Rabiee, Ella Rabinovich, Karthik Radhakrishnan, Bhaktipriya Radharapu, Kira Radinsky, Taki Hasan Rafi, Davood Rafiei, Alessandro Raganato, Luca Ragazzi, Aditi Raghunathan, Vethavikashini Chithrra Raghuram, Tathagata Raha, Razieh Rahimi, Mizanur Rahman, Hossein A. Rahmani, Daking Rai, Nishat Raihan, Roberta Raileanu, Lisa Raithel, Bhiksha Raj, Chahat Raj, Sahar Rajabi, Rishanth Rajendhran, Sello Ralethe, Ganesh Ramakrishnan, Karthikeyan Natesan Ramamurthy, Bhaskar Ramasubramanian, Giulia Rambelli, Owen Rambow, Aida Ramezani, Roshni Ramnani, Sahana Ramnath, Rita Parada Ramos, Sanjana Ramprasad, Nitin Ramrakhiyani, Pritika Ramu, Guillem Ramírez, Yide Ran, Federico Ranaldi, Huzefa Rangwala, Dongning Rao, Gengyu Rao, Jinmeng Rao, Jun Rao, Farzana Rashid, Royi Rassin, Vipul Kumar Rathore, Shivam Raval, Sahithya Ravi, Sujith Ravi, Tejasvi Ravi, Abhilasha Ravichander, Ajay Madhavan Ravichandran, Baishakhi Ray, Pretam Ray, Gabor Recski, Avinash Reddy, Revanth Gangi Reddy, Varshini Reddy, Emily Reif, Yuval Reif, Nils Reimers, Paul Reisert, Nils Reiter, Sravan Reddy Rekula, Feiliang Ren, Han Ren, Houxing Ren, Jie Ren, Kan Ren, Kui Ren, Lei Ren, Pengjie Ren, Runtao Ren, Shuo Ren, Weijieying Ren, Xiaozhe Ren, Xubin Ren, Yupei Ren, Yarik Menchaca Resendiz, Fabian Retkowski, MohammadHossein Rezaei, Ryokan Ri, Eugénio Ribeiro, Kathryn Ricci, Enora Rice, Stephen D. Richardson, Korbinian Riedhammer, Jason Riesa, Maarten de Rijke, Parker Riley, Laura Rimell, Annette Rios, Kumar Rishabh, Yong Man Ro, Gabriel Roccabruna, Gil Rocha, Julian Martin Rodemann, Álvaro Rodrigo, Christophe Rodrigues, Miguel E. Rodríguez, Melissa Roemmele, Mark Rofin, Timothy T. Rogers, Valeriya Rogovchenko, Lina Maria Rojas-Barahona, Nathan Roll, Muhammad Roman, Srikanth Ronanki, Wenge Rong, Subendhu Rongali, Zhu Ronghang, Weng Rongxiang, Robert Van Rooij, Amirmohammad Rooshenas, Tanya Roosta, Christophe Ropers, Carolyn Rose, Jonathan Rose, Shachar Rosenman, Alexis Ross, Robert Ross, Mohammad Rostami, Aiala Rosá, Sascha Rothe, Dmitri Roussinov, Dimitris Roussis, Bryan R. Routledge, Aniruddha Roy, Aurko Roy, Deb Roy, Shamik Roy, Soumyadeep Roy, Sudipta Singha Roy, Sumegh Roychowdhury, Aswin Rrv, Jiacheng Ruan, Jie Ruan, Jingqing Ruan, Tong Ruan, Weitong Ruan, Terry Ruas, Patrick Ruch, Rachel Ruder, William Rudman, Ewa Rudnicka, Alexander Rudnick, Koustav Rudra, Daniel Ruffinelli, Li Rui, Mukund Rungta, Tuukka Ruotsalo, Benjamin Matthias Ruppik, Jonathan Rusert, Stefan Ruseti, Jenna Russell, Daniel Russo, Irene Russo, Phillip Rust, Derek Ruths, Elena Sofia Ruzzetti, Max Ryabinin, Hale Sirin Ryan, Michael J Ryan, Pavel Rychlý, Sangwon Ryu, Susanna Rücker, Sara rajae

Ramaneswaran S, Shwetha S, Jon Saad-Falcon, Ashish Sabharwal, Ahmed Sabir, Sahand Sabour, Rachneet Singh Sachdeva, Kaushik Ram Sadagopan, Abdelrahman Boda Sadallah, Farig Sadeque, Arka Sadhu, Philipp Sadler, Amir Saeidi, Kate Saenko, Abdulfattah Safa, Mustafa Safdari, Benoît Sagot, Amit Kumar Sah, Punyajoy Saha, Rohan Saha, Sougata Saha, Sriparna Saha, Sattvik Sahai, Vaibhav Sahai, Saurav Sahay, Magnus Sahlgren, Dhruv Sahnun, Nihar Ranjan Sahoo, Oscar Sainz, Itsumi Saito, Kuniko Saito, Yusuke Sakai, Hiroki Sakaji, Andrey Sakhovskiy, Alireza Salemi, Jonne Saleva, Vishal Vivek Saley, Suchir Salhan, Flora Salim, Mikhail Salnikov, Tanja Samardzic, Niloofer Safi Samghabadi, Farhan Samir, Ranbir Singh Sanasam, Kate Sanders, Jivnesh Sandhan, Raghav Sandhu, Erik F. Tjong Kim Sang, Jitao Sang, Manuela Sanguinetti, Lorenzo Sani, Ashwin

Sankar, Scott Sanner, Brenda Salenave Santana, Bishal Santra, Debarshi Kumar Sanyal, Soumya Sanyal, Xabier Saralegi, Vasanth Sarathy, Sunita Sarawagi, Gabriel Herbert Sarch, Ruhi Sarikaya, Chayan Sarkar, Rupak Sarkar, Souvika Sarkar, Sujoy Sarkar, Ritesh Sarkhel, Shota Sasaki, David Sasu, Ranjan Satapathy, Shiki Sato, Shin'ichi Satoh, Prasanna Sattigeri, Andrey Savchenko, Beatrice Savoldi, Guergana K Savova, Tomohiro Sawada, Albert Sawczyn, Rohit Saxena, Michael Saxon, Daniel Scalena, Alexander Scarlatos, Elizabeth J. Schaefer, Emmanuel Schang, Markus Schedl, Yves Scherrer, Matthias Scheutz, Benjamin Schiller, David Schlangen, Patrícia Schmidová, Vera Schmitt, Tobias Schnabel, Gerold Schneider, Nathan Schneider, Stephanie Schoch, Timo Pierre Schrader, Patrick Schramowski, Christopher Schröder, Dan Schumacher, Jakob Schuster, Ulrich Schäfer, Wesley S Scivetti, Ivan Sedykh, Simon See, Ethan Seefried, Christin Seifert, Amit Seker, Toshiyuki Sekiya, Bilgehan Sel, David Semedo, Sina Semnani, Jaydeep Sen, Lütfi Kerem Senel, Ayan Sengupta, Meghdut Sengupta, Sailik Sengupta, Jaehyung Seo, Jeongyeon Seo, Minhyuk Seo, Paul Hongsuck Seo, Dawn Sepehr, Giuseppe Serra, Sofia Serrano, Ashish Seth, Skyler Seto, Alay Dilipbhai Shah, Darsh Jaidip Shah, Mubarak Shah, Raj Sanjay Shah, Shalin Shah, Uri Shaham, Haz Sameen Shahgir, Rifat Shahriyar, Rana Shahroz, Chantal Shaib, Valerie Shalin, Mehrnoush Shamsfard, Hengchao Shang, Lanyu Shang, Lifeng Shang, Lin Shang, Mingyue Shang, Chen Shani, Aditya Shanmugham, Jian Shao, Jing Shao, Minglai Shao, Ruichen Shao, Rulin Shao, Wei Shao, Yanjun Shao, Yihua Shao, Yijia Shao, Yunfan Shao, Zirui Shao, Natalie Shapira, Omar Sharif, Abhishek Sharma, Akshat Sharma, Amit Sharma, Divya V Sharma, Harsh Sharma, Karishma Sharma, Kartik Sharma, Rajesh Sharma, Serge Sharoff, Shuaijie She, Ryan Shea, Shady Shehata, Sumit Shekhar, Artem Shelmanov, Chao Shen, Ethan Shen, Guangyu Shen, Hua Shen, Huawei Shen, Jiayi Shen, Jocelyn J Shen, Kai Shen, Lei Shen, Li Shen, Linlin Shen, Ming Shen, Sherrie Shen, Siqi Shen, Tao Shen, Tianhao Shen, Weizhou Shen, Xiangqing Shen, Xinyue Shen, Xu Shen, Yaling Shen, Yanzhen Shen, Ying Shen, Yinghan Shen, Zejiang Shen, Zhan Shen, Jiawei Sheng, Qiang Sheng, Quan Z. Sheng, Tao Sheng, Zhang Jia Sheng, Shaul Sheinhav, Amit Sheth, Ivaxi Sheth, Anudeex Shetty, Pranav Shetty, Bowen Shi, Chuan Shi, Chufan Shi, Dan Shi, Daqian Shi, Haizhou Shi, Han Shi, Haochen Shi, Haoxiang Shi, Jie Shi, Kaize Shi, Kejian Shi, Lida Shi, Ning Shi, Qi Shi, Tianyuan Shi, Tongyue Shi, Weijia Shi, Wenqi Shi, Wentao Shi, Xiao Shi, Xiaoming Shi, Yucheng Shi, Zhengliang Shi, Zijing Shi, Tomohide Shibata, G. M. Shahariar Shibli, Min-Han Shih, Li Shilong, Shuichiro Shimizu, Hidetoshi Shimodaira, Anastasia Shimorina, Haebin Shin, Hyopil Shin, Jamin Shin, Jinwoo Shin, Jong-hun Shin, Saim Shin, Youhyun Shin, Kazutoshi Shinoda, Kiyoaki Shirai, Yow-Ting Shiue, Namrata Shivagunde, Michal M Shlapentokh-Rothman, Michal Shmueli-Scheuer, Eyal Shnarch, Mohammad Shoebybi, Reza Shokri, Lidan Shou, Zheng Shou, Ingroj Shrestha, Dong Shu, Kai Shu, Yiheng Shu, Hong-Han Shuai, Liu Shudong, Wang Shuhe, Ruihao Shui, Zeren Shui, Shreya Shukla, KaShun Shum, Li Shuo, Ammon C Shurtz, Alexander Shvets, Mei Si, Shijing Si, Shuzheng Si, Anthony Sicilia, Lucia Siciliani, Nikhil Siddhartha, A.B. Siddique, Grigori Sidorov, Noah Yamamoto Siegel, Sandipan Sikdar, Nisansa de Silva, Patrick Queiroz Da Silva, Mong Yuan Sim, Michel Simard, Jakub Simko, Marian Simko, Arabella Jane Sinclair, Apoorva Singh, Arshdeep Singh, Avyav Kumar Singh, Bhavyajeet Singh, Joykirat Singh, Mantek Singh, Mayank Singh, Mukul Singh, Muskaan Singh, Prateek Singh, Rituraj Singh, Sagalpreet Singh, Sameer Singh, Shivalika Singh, Siffi Singh, Mohit Singhal, Karan Singla, Parag Singla, Iustin Sirbu, Clemencia Siro, Kairit Sirts, Anushka Sivakumar, Jassivan Alex Sivakumar, Steven Skiena, Paweł Marek Skórzewski, Diego Fernandez Slezak, Aviv Slobodkin, Charese Smiley, David A. Smith, Tony C Smith, Pavel Smrz, Răzvan-Alexandru Smădu, Vésteinn Snæbjarnarson, Shagun Sodhani, Sokratis Sofianopoulos, Jy-yong Sohn, Mohammad Golam Sohrab, Petr Sojka, Sai Ashish Somayajula, Pia Sommerauer, Guijin Son, Junyoung Son, Seungwoo Son, Suhyune Son, Dandan Song, Dawei Song, Dawn Song, Demin Song, Guojie Song, Huacheng Song, Hwanjun Song, Hyun-Je Song, Jongyoon Song, Kai Song, Kaisong Song, Kaitao Song, Kyungwoo Song, Lingyun Song, Linqi Song, Linxin Song, Maojia Song, Mingyang Song, Peiyang Song, Ruihua Song, Siyuan Song, Xia Song, Xiaoying Song, Xin Song, Xinyuan Song, Yang Song, Yewei Song, Yifan Song, Yingjin Song, Yiping Song, Yixiao Song, Yuanfeng

Song, Yueqi Song, Yufan Song, Zhao Song, Zirui Song, Nikita Soni, Sandeep Soni, Shashank Sonkar, Ekta Sood, Lay-Ki Soon, Alessandro Sordoni, Jeffrey S. Sorensen, Ionut Sorodoc, Alexey Sorokin, Nikita Sorokin, Anna Sotnikova, Zhivar Sourati, Rui Sousa-Silva, Alexander Spangher, Lucia Specia, Matthias Sperber, Evangelia Spiliopoulou, Andreas Spitz, Maximilian Spliethöver, Zayne Rea Sprague, Makesh Narsimhan Sreedhar, Rohini Srihari, Neha Srikanth, Krishna Srinivasan, Prashanth Srinivasan, Tejas Srinivasan, Vijay Srinivasan, Somayajulu Sripada, Aarozi Srivastava, Aseem Srivastava, Saurabh Srivastava, Kv Aditya Srivatsa, Steffen Staab, Joe Stacey, Dominik Stammbach, Karolina Stanczak, David Stap, Katherine Stasaski, Mark Steedman, Sjoerd van Steenkiste, Nicolas Stefanovitch, Sebastian Steindl, Hayk Stepanyan, Andreas Stephan, Igor Sterner, Hannah Sterz, Julius Steuer, Regina Stodden, Alessandro Stolfo, Shane Storks, Svetlana Stoyanchev, Ivelina Stoyanova, Markus Strohmaier, Sebastian Stüker, Hang Su, Hsuan Su, Hui Su, Hung-Ting Su, PuZhen Su, Sen Su, Weihang Su, Xiangdong Su, Xin Su, Yi Su, Ying Su, Zhaochen Su, Zhe Su, Prasanna Lakkur Subramanyam, Katsuhito Sudoh, Sainbayar Sukhbaatar, Oren Sultan, Changlong Sun, Chengjie Sun, Chenkai Sun, Chenxi Sun, Dachun Sun, Guangzhi Sun, Guangzhong Sun, Hanshi Sun, Hao Sun, Haotian Sun, Hongda Sun, Jiacheng Sun, Jingyi Sun, Kaiser Sun, Kun Sun, Le Sun, Linzhuang Sun, Maosong Sun, Penglei Sun, Qiao Sun, Qing Sun, Qiushi Sun, Rui Sun, Ruoyu Sun, Sann Sun, Shiliang Sun, Shuo Sun, Si Sun, Tiening Sun, Wei Sun, Xiaofei Sun, Xiaoshuai Sun, Xin Sun, Xing Sun, Xingpeng Sun, Xu Sun, Yang Sun, Yingfei Sun, Yizheng Sun, Yizhou Sun, Youcheng Sun, Yuanyuan Sun, Yuqing Sun, Yuxi Sun, Zengkui Sun, Zequn Sun, Zhaoyue Sun, Zhewei Sun, ZhongXiang Sun, Zhouhao Sun, Jayaprakash Sundararaj, Dhanasekar Sundararaman, Megha Sundriyal, Hakyung Sung, Mujeen Sung, Yoo Yeon Sung, Yun-Hsuan Sung, Marek Suppa, Siddharth Suresh, Aditya Kaushik Surikuchi, Lintang Sutawika, Ashima Suvarna, Masahiro Suzuki, Anej Svete, Jakub Szymanik, Ricardo Muñoz Sánchez, Jyoti Kunal shah, Yurun song

Jacek Tabor, Hideyuki Tachibana, Jaesung Tae, Andrea Tagarelli, Chihiro Taguchi, Nina Tahmasebi, Hagai Taitelbaum, Dima Taji, Yu Takagi, Koh Takeuchi, Koichi Takeuchi, Tetsuya Takiguchi, Ekaterina Taktasheva, Indraayudh Talukdar, Partha Talukdar, Ankur Taly, Kar Yan Tam, Yik-Cheung Tam, Srikanth G. Tamilselvam, Chao-Hong Tan, Cheng Tan, Chuanyuan Tan, Fei Tan, Hanzhuo Tan, Haochen Tan, Hongye Tan, Qiaoyu Tan, Shaomu Tan, Shicheng Tan, Tao Tan, Tian Tan, Weiting Tan, Xiaoyu Tan, Xin Tan, Yi Chern Tan, Yue Tan, Zhen Tan, Zhen-tao Tan, Hideki Tanaka, Buzhou Tang, ChiaWei Tang, Duyu Tang, Haoran Tang, Jie Tang, Jing Tang, Kewen Tang, Liyan Tang, Nan Tang, Qiaoyu Tang, Ruiming Tang, Ruixiang Tang, Xiaohang Tang, Xuemei Tang, Xunzhu Tang, Yihong Tang, Yixuan Tang, Yu-Chien Tang, Yubao Tang, Yun Tang, Zheng Tang, Zhenheng Tang, Zhiwen Tang, Zichen Tang, Motoki Taniguchi, Joshua Tanner, Eshaan Tanwar, Chaofan Tao, Chenyang Tao, Dacheng Tao, Dehao Tao, Jianhua Tao, Wenbiao Tao, Zhengwei Tao, Zhiqiang Tao, Zhuo Tao, Maksym Taranukhin, Panuthep Tasawong, Tarun Tater, Michiaki Tatsubori, Christopher Tauchmann, Ateret Anaby Tavor, Ran Tavory, Anuja Tayal, Matthieu Tehenan, Selma Tekir, Maria Teleki, Atharva Tendle, Joshua B. Tenenbaum, Chong Teng, Joel R. Tetreault, Nithum Thain, Megh Thakkar, Nandan Thakur, Ramya Keerthy Thatikonda, Barry-John Theobald, Krishnaprasad Thirunarayan, Craig Alexander Thomson, Camilo John Thorne, James Thorne, Veronika Thost, Poojitha Thota, Ye Kyaw Thu, Bozhong Tian, Jidong Tian, Jinchuan Tian, Junfeng Tian, Lin Tian, Shiyu Tian, Yu Tian, Yuan Tian, Yuandong Tian, Yuchen Tian, Yufei Tian, Jörg Tiedemann, Roelien C Timmer, Tsang Hong Ting, Abhisek Tiwari, Prayag Tiwari, Vashisth Tiwari, Lindia Tjuatja, Assaf Toledo, Katrin Tomanek, Nicholas Tomlin, Antonela Tommasel, Hanghang Tong, MeiHan Tong, Meng Tong, Minghao Tong, Jonathan Tonglet, Francesca Toni, Atnafu Lambebo Tonja, Manuel Tonneau, Shakila Mahjabin Tonni, Cagri Toraman, Tiago Timponi Torrent, Rishit Toteja, Samia Touileb, Hawau Olamide Toyin, Masashi Toyoda, Hieu Tran, Khiem Vinh Tran, Quyen Tran, Son Quoc Tran, Thai Dang Tran, Thanh Tran, Isabel Trancoso, David Traum, Sony Trenous, Volker Tresp, Ilias Triantafyllopoulos, Harsh Trivedi, Shubhendu Trivedi, Sean Trott, Yun Da Tsai, Dimitrios Tsarapatsanis, Bo-Hsiang Tseng,

Masaaki Tsuchida, Hayato Tsukagoshi, Yuiko Tsunomori, Akim Tsvigun, Geng Tu, Lifu Tu, Mei Tu, Shangqing Tu, Zhiying Tu, Zhuowen Tu, Hessel Tuinhof, Elena Tutubalina, Aman Tyagi, Ut-karsh Tyagi, Georgios Tzimiropoulos, Zhang wen tao

Akhil Udathu, Md Nayem Uddin, Can Udomcharoenchaikit, Ryo Ueda, Gökçe Uludoğan, Inigo Jauregi Unanue, Lyle Ungar, Rheeya Uppaal, Gorka Urbizu, Ashok Urlana, Asahi Ushio, Nicolas Usunier, David Uthus, Takehito Utsuro

Kavin R V, Venktesh V, Robert Vacareanu, Ashwini Vaidya, Andric Valdez Valenzuela, Mojtaba Valipour, Estrella Vallecillo-Rodríguez, Sai P Vallurupalli, Linh Ngo Van, Vasudha Varadarajan, Praveen Srinivasa Varadhan, Francielle Vargas, Amey Varhade, Siddharth Varia, Vasudeva Varma, Lav R. Varshney, Prasoon Varshney, Rossella Varvara, Roman Vashurin, Oleg Vasilyev, Justin Vasselli, Rosni Vasu, Peerapon Vateekul, Pedro Vaz-de-Melo, Artem Vazhentsev, Michalis Vazirgiannis, Eva Maria Vecchi, Hariram Veeramani, Dan John Velasco, Julien Velcin, Sivapriya Vellaichamy, Raviteja Vemulapalli, Suresh Venkatasubramanian, Kavana Venkatesh, Thejas Venkatesh, Praveen Venkateswaran, Tessa Verhoef, Apurv Verma, Arun Verma, Neha Verma, Nikhil Verma, Rakesh Verma, Gautier Viaud, Peter Vickers, Gianluca Vico, Bertie Vidgen, Renata Vieira, Lovekesh Vig, Ashwin Kalyan Vijayakumar, Anvesh Rao Vijjini, David Vilares, Dan Vilenchik, Danae Sanchez Villegas, Vibhav Vineet, Bhanukiran Vinzamuri, Koushik Viswanadha, Amar Viswanathan, Nikolas Vitsakis, Jean-Noël Vittaut, Andreas Vlachos, Piek Vossen, Pavlos Vougiouklis, Trang Vu, Slobodan Vucetic, Renato Vukovic, Ugurcan Vurgun, Abhijeet Vyas, Ivan Vykopal, Ekaterina Vylomova, Deeksha varshney

Araloak W, Takashi Wada, Rohan Wadhawan, Eitan Wagner, Jan Philip Wahle, Hiromi Wakaki, Shoko Wakamiya, Andreas Waldis, Jonas Wallat, David Wan, Guancheng Wan, Hai Wan, Herun Wan, Mengting Wan, Stephen Wan, Yao Wan, Yingjia Alisa Wan, Yu Wan, Zhen Wan, Zhongwei Wan, An Wang, Bang Wang, Baoxin Wang, BiChen Wang, Bicong Wang, Bin Wang, Bing Wang, Bingning Wang, Bo Wang, Boxiang Wang, Chaojie Wang, Chaokun Wang, Chen Wang, Chenglong Wang, Chenguang Wang, Chenhao Wang, Chenxi Wang, Chenxu Wang, Chuan-Ju Wang, Cong Wang, Cunxiang Wang, Daling Wang, Danding Wang, Darren Yow-Bang Wang, Di Wang, Dingmin Wang, Dingzirui Wang, Fali Wang, Futing Wang, Guanchu Wang, Hai Wang, Haifeng Wang, Haishuai Wang, Hanqing Wang, Hao Wang, Haobo Wang, Haochun Wang, Haodong Wang, Haohan Wang, Haoran Wang, Haoyu Wang, Heng Wang, Hengyi Wang, Hongru Wang, Hongya Wang, Houfeng Wang, Hsin Wei Wang, Hsin-Min Wang, Huan Wang, Hui Wang, Huimin Wang, Huimu Wang, Jia-ning Wang, Jiakai Wang, Jian Wang, Jianxin Wang, Jianyong Wang, Jianyou Wang, Jianzong Wang, Jiapeng Wang, Jiashuo Wang, Jiawei Wang, Jiayi Wang, Jie Wang, Jingang Wang, Jingquan Wang, Jingwei Wang, Jingyu Wang, Jinpeng Wang, Jinqiao Wang, Jun Wang, Junda Wang, Junjie Wang, Junli Wang, Junyang Wang, Junzhe Wang, Kai Wang, Kangrui Wang, Kexin Wang, Lei Wang, Li Wang, Liang Wang, Lianxi Wang, Lijuan Wang, Limin Wang, Lin Wang, Linlin Wang, Lu Wang, Mengru Wang, Mengyu Wang, Ming Wang, Minghan Wang, Mingyang Wang, Minzheng Wang, Pancheng Wang, Pei Wang, Peiqi Wang, Peiran Wang, Peiyi Wang, Pengyu Wang, Ping Wang, Pingjie Wang, Qianli Wang, Qianlong Wang, Qidong Wang, Qihao Wang, Qingni Wang, Qiqi Wang, Qiufeng Wang, Quan Wang, Renxi Wang, Rui Wang, Ruibo Wang, Ruida Wang, Runhui Wang, Ruoyao Wang, Shangfei Wang, Shaobo Wang, Shaojun Wang, Sheng Wang, Shi Wang, Shiqi Wang, Shuaiqiang Wang, Shuang Wang, Shun Wang, Shuo Wang, Shuting Wang, Siqi Wang, Siyin Wang, Siyuan Wang, Song Wang, Suge Wang, Suhang Wang, Teng Wang, Tengjiao Wang, Tianchun Wang, Tianduo Wang, Tianshu Wang, Ting Wang, Wei Wang, WeiYun Wang, Weichuan Wang, Weihua Wang, Weixuan Wang, Weizhi Wang, Wen Wang, Wenjin Wang, Wenwu Wang, Wenxuan Wang, Xiang Wang, Xiangfeng Wang, Xiao Wang, Xiaobo Wang, Xiaochen Wang, Xiaohua Wang, Xiaojie Wang, Xiaoli Wang, Xiaoling Wang, Xiaolong Wang, Xiaoman Wang, Xiaoxuan Wang, Xiaoyang Wang, Xiaxia Wang,



Xidong Wang, Xin Wang, Xinghao Wang, Xinglin Wang, Xinpeng Wang, Xintao Wang, Xintong Wang, Xinyou Wang, Xinyu Wang, Xinze Wang, Xinzuo Wang, Xu Wang, Xuehai Wang, Xuepeng Wang, Xueqian Wang, Xuwu Wang, Yadao Wang, Yajing Wang, Yanbo Wang, Yanhao Wang, Yanlin Wang, Yanling Wang, Yaoming Wang, Yaowei Wang, Yaqing Wang, Yasha Wang, Ye Wang, Yequan Wang, Yi Wang, Yibo Wang, Yichao Wang, Yichen Wang, Yifan Wang, Yifei Wang, Yike Wang, Yilei Wang, Yiming Wang, Yimu Wang, Ying Wang, Yinggui Wang, Yinghan Wang, Yinglin Wang, Yining Wang, Yiran Wang, Yixuan Wang, Yongjie Wang, Yu Wang, Yuanzhuo Wang, Yuchi Wang, Yudong Wang, Yue Wang, Yueqian Wang, Yufei Wang, Yuhao Wang, Yujie Wang, Yumeng Wang, Yuqi Wang, Yuxia Wang, Yuxin Wang, Yuxuan Wang, Zecheng Wang, Zehan Wang, Zekun Wang, Zezhong Wang, Zhaoyang Wang, Zhefeng Wang, Zhen Wang, Zheng Wang, Zhenglin Wang, Zhengyang Wang, Zhenhailong Wang, Zhenting Wang, ZhiZheng Wang, Zhihao Wang, Zhixiang Wang, Zhiyong Wang, Zhiyuan Wang, Zhongyuan Wang, Zhuoer Wang, Zichao Wang, Zifeng Wang, Zihan Wang, Ziheng Wang, Zijie Wang, Zili Wang, Zilong Wang, Zimu Wang, Ziqi Wang, Ziyang Wang, Ziyue Wang, Zun Wang, Dittaya Wanvarie, Azmine Touseh Wasi, Julia Watson, Ishaan Watts, Lucas Weber, Ruwan Weerasinghe, Bowen Wei, Chen Wei, Daimeng Wei, Furu Wei, Hao-Ran Wei, Haoran Wei, Hua Wei, Jingxuan Wei, Kaiwen Wei, Kangda Wei, Lingwei Wei, Penghui Wei, Sheng-Lun Wei, Wei Wei, Xiao Wei, Xiaochi Wei, Xiaoyong Wei, Ying Wei, Zhepei Wei, Zhihua Wei, Zhongyu Wei, Zihao Wei, Maxwell Weinzierl, Gail Weiss, Fei Weizhi, Charles Welch, Chris Welty, Bingbing Wen, Juan Wen, Liang Wen, Xiaofei Wen, Ximing Wen, Yanlong Wen, Zhen Wen, Zhihao Wen, Zhizhao Wen, James Bradley Wendt, Fenghua Weng, Tsui-Wei Weng, Yixuan Weng, Jason E Weston, Tillman Weyde, Joyce Jiyoung Whang, Haryo Akbarianto Wibowo, Rachel Wicks, Michael Wiegand, Adam Wiemerslage, Adam Wierzbicki, Matthew Wiesner, Ethan Wilcox, Matthew Wilkens, Rodrigo Wilkens, Dane A Williamson, Steven R. Wilson, David Wingate, Michael J. Witbrock, Sondre Wold, Tomer Wolfson, Eric Wong, Ginny Wong, Ngai Wong, Wichayaporn Wongkamjan, Zach Wood-Doughty, Simon Woodhead, Phil Woodland, Anna Wróblewska, Alexander Wu, Anne Wu, Ben Peng Wu, Bin Wu, Bohao Wu, Bowen Wu, Dayong Wu, Dekun Wu, Di Wu, Fang Wu, Feijie Wu, Han Wu, Haolun Wu, Haoning Wu, Haoyi Wu, Haoyuan Wu, Honghan Wu, Hui Wu, Jian Wu, Jiancan Wu, Jiayi Wu, Jie Wu, Jincenzi Wu, Jinge Wu, Junda Wu, Junhong Wu, Junkang Wu, Junshuang Wu, Kangxi Wu, Lianwei Wu, Likang Wu, Linjuan Wu, Minghao Wu, Mingyan Wu, Muling Wu, Peilin Wu, Peng Wu, Qianhui Wu, Qinzhuo Wu, Sai Wu, Shangyu Wu, Shengguang Wu, Shu Wu, Sifan Wu, Siwei Wu, Sixing Wu, Taiqiang Wu, Ting-Wei Wu, Tongtong Wu, Wei Wu, Weiqi Wu, Weiyi Wu, Wenhao Wu, Winston Wu, Xiaoting Wu, Xin Wu, Xinbo Wu, Xindong Wu, Xinwei Wu, Xixin Wu, Xiyang Wu, Xuansheng Wu, Xuecheng Wu, Yang Wu, Yangjun Wu, Yifan Wu, Yihan Wu, Yingnian Wu, Yixin Wu, Yu Wu, Yuan Wu, Yuanhao Wu, Yuchuan Wu, Yuhao Wu, Yulong Wu, Yuping Wu, Yurong Wu, Zekun Wu, Zengqing Wu, Zhaofeng Wu, Zhaoxuan Wu, Zhengxuan Wu, Zhenyu Wu, Zhihao Wu, Zhijing Wu, Zihui Wu, Zijun Wu, Zongru Wu, Zongyu Wu, Xie Jia wen

Alexandros Xenos, Wang Xi, Zhaoan Xi, Zhiheng Xi, Zhou Xi, Feng Xia, Haotian Xia, Lei Xia, Lianghao Xia, Peng Xia, Qingrong Xia, Rui Xia, Tian Xia, Tingyu Xia, Wei Xia, Xiao Xia, Yan Xia, Yu Xia, Yuxi Xia, Hu Xiang, Lu Xiang, Xi Xiangyu, Chaojun Xiao, Chenghao Xiao, Chuan Xiao, Chunyang Xiao, Jin Xiao, Jinfeng Xiao, Jun Xiao, Meng Xiao, Qingfa Xiao, Teng Xiao, Yao Xiao, Yijia Xiao, Yisheng Xiao, Yunxuan Xiao, Yunze Xiao, Yuxin Xiao, Zeguan Xiao, Zhiqing Xiao, Zilin Xiao, Li Xiaoguang, Chen Xiaoshu, Dong Xiaoyu, Alex Xie, Binzhu Xie, Chenhao Xie, Haoran Xie, Hong Xie, Hongyan Xie, Kaige Xie, Lei Xie, Pengjun Xie, Ruobing Xie, Sean Xie, Tianbao Xie, Wenya Xie, Xiaofei Xie, Yong Xie, Yuanzhen Xie, Yubo Xie, Yueqi Xie, Yunhe Xie, Yuqiang Xie, Yuxi Xie, Zhixin Xie, Zhongbin Xie, Zhuohan Xie, Amy Xin, Chunlei Xin, Liao Xinfeng, Bowen Xing, Chen Xing, Hangdi Xing, Linzi Xing, Long Xing, Rui Xing, Shuo Xing, Xingrun Xing, Boya Xiong, Caiming Xiong, Feiyu Xiong, Haoyi Xiong, Hui Xiong, Jing Xiong, Jinjun Xiong, Kai Xiong, Siheng Xiong, Wei Xiong, Weimin Xiong, Ying Xiong, Yizhe Xiong, Yun Xiong, Wei Wang Xjtlu, Baixuan Xu, Benfeng Xu, Bing Xu, Bo Xu, Boyan Xu, Cai

Xu, Canwen Xu, Chejian Xu, Chen Xu, Cheng-zhong Xu, Dehong Xu, Depeng Xu, Derong Xu, Dongkuan Xu, Fangzhi Xu, Guandong Xu, Hainiu Xu, Haiyang Xu, Han Xu, Hanzi Xu, Hao Xu, Haotian Xu, Heng-Da Xu, Hong Xu, Hongling Xu, Hongshen Xu, Hongxia Xu, Hongzhi Xu, Hua Xu, Jia Xu, Jiahao Xu, Jing Xu, Jingjing Xu, Jun Xu, Jundong Xu, Justin Xu, Kaidi Xu, Kaishuai Xu, Kun Xu, Lei Xu, Liyan Xu, Mayi Xu, Ming Xu, Nan Xu, Paiheng Xu, Pengyu Xu, Ran Xu, Rongwu Xu, Rui Xu, Ruochen Xu, Ruoxi Xu, Shanshan Xu, Tianyang Xu, Ting Xu, Wang Xu, Weijia Xu, Weikai Xu, Weiran Xu, Weiwen Xu, Xiangmin Xu, Xiangzhe Xu, Xin Xu, Xinrun Xu, Xiong Xiao Xu, Xueqiang Xu, Yan Xu, Yang Xu, Yangyifan Xu, Yige Xu, Yongxin Xu, Yongxiu Xu, Yunqiu Xu, Zenan Xu, Zhangchen Xu, Zhao Xu, Zhe Xu, Zhen Xu, Zheng Xu, Zhenran Xu, Zhichao Xu, Zhijian Xu, Zhikun Xu, Zhipeng Xu, Zhiyang Xu, Zihan Xu, Zihang Xu, Ziyao Xu, Zunnan Xu, Luo Xuan, Qi Xuan, Richeng Xuan, Boyang Xue, Huiyin Xue, Jiaqi Xue, Mingfeng Xue, Siqiao Xue, Wei Xue, Yun Xue, Zhiyu Xue, Christos Xypolopoulos

Ajay Yadav, Neemesh Yadav, Srishti Yadav, Vikas Yadav, Aditya Yadavalli, Kosuke Yamada, Hiroaki Yamagiwa, Atsuki Yamaguchi, Takateru Yamakoshi, Michiharu Yamashita, Bi Cheng Yan, Bing Yan, Cheng Yan, Dong Yan, Junbing Yan, Junchi Yan, Le Yan, Ming Yan, Ruiyi Yan, Shaotian Yan, Shi-Qi Yan, Shuicheng Yan, Shuo Yan, Xiaohui Yan, Xifeng Yan, Xin Yan, Yang Yan, Yibo Yan, Yukun Yan, Aimin Yang, An Yang, Baosong Yang, Bohao Yang, Boyang Yang, Carl Yang, Cehao Yang, Chang Yang, Changbing Yang, Chao Yang, Chaoqun Yang, Cheng Yang, Chenghao Yang, Chenxu Yang, Chih-Kai Yang, Dingkang Yang, Dingyi Yang, Dong Yang, Fan Yang, Fangkai Yang, Fengyi Yang, Guang Yang, Guanqun Yang, Hao Yang, Haoran Yang, Heng Yang, Hongxia Yang, Hou Yang, Huichen Yang, Ivory Yang, Jack Yang, Jian Yang, Jiayi Yang, Jinfa Yang, Jun Yang, Kaicheng Yang, Kailai Yang, Kejuan Yang, Kexin Yang, Li Yang, Liang Yang, Liangwei Yang, Libin Yang, Liner Yang, Longfei Yang, Ming-Hsuan Yang, Mu Yang, Nakyeong Yang, Nan Yang, Pei Yang, Qing Yang, Rui Yang, Ruixin Yang, Ruosong Yang, Sen Yang, Shiping Yang, Shu Yang, Shuo Yang, Sohee Yang, Song Yang, Soyoung Yang, Tao Yang, Tiankai Yang, Tianyu Yang, Wanli Yang, Wen Yang, Wenmian Yang, Xiao-wen Yang, Xiaocheng Yang, Xiaoda Yang, Xincheng Yang, Xinyi Yang, Xinyu Yang, Xiulin Yang, Xuwen Yang, Yahan Yang, Yajing Yang, Yaming Yang, Yao Yang, Yaodong Yang, Yating Yang, Yi Yang, Yifan Yang, Yifei Yang, Yiming Yang, Yiyuan Yang, Yongjin Yang, Yuanhang Yang, Yue Yang, Yuhao Yang, Yujiu Yang, Yuming Yang, Yuqi Yang, Yuqing Yang, Ze Yang, Zeyuan Yang, Zhe Yang, Zhengyuan Yang, Zhengzhe Yang, Zhichao Yang, Zhicheng Yang, Zhiyu Yang, Zhou Yang, Zhuoyi Yang, Zi Yang, Ziqing Yang, Roman Yangarber, Gong Yangrui, Ken Yano, Chang Yao, Jianmin Yao, Jiarui Yao, Jiashu Yao, Jihan Yao, Jing Yao, Kaichun Yao, Liang Yao, Lili Yao, Lina Yao, Maxwell Yao, Peiran Yao, Quanming Yao, Wei Yao, Xingcheng Yao, Yao Yao, Yixiang Yao, Yuan Yao, Yuekun Yao, Yuhang Yao, Zhewei Yao, Zijun Yao, Mahsa Yarmohammadi, Semih Yavuz, Mert Yazan, Bingyang Ye, Deheng Ye, Haoran Ye, Haotian Ye, Jiabo Ye, Jieping Ye, Jingheng Ye, Junjie Ye, Rong Ye, Wei Ye, Wentao Ye, Xiao Ye, Yixin Ye, Yuyang Ye, Anar Yeginbergen, Eric Yeh, Sung-Lin Yeh, Asaf Yehudai, Seren Yenikent, Eunjung Yeo, Jeonghun Yeo, Meliha Yetisgen, Bowen Yi, Euiin Yi, Mun Yong Yi, Siyu Yi, Xiaoyuan Yi, Xue Yida, Emine Yilmaz, Jinyeong Yim, Wen-wai Yim, Bing Yin, Congchi Yin, Dawei Yin, Heqi Yin, Hongzhi Yin, Jianwei Yin, Lu Yin, Wenbiao Yin, Wotao Yin, Xiang Yin, Xunjian Yin, Yichun Yin, Yu Yin, Yuwei Yin, Zhangyue Yin, Zhenfei Yin, Ziyi Yin, Gal Yona, Wang Yong, Zheng Xin Yong, Eunseop Yoon, Jaehong Yoon, Seunghyun Yoon, Sungroh Yoon, Sunjae Yoon, WonJin Yoon, Ori Yoran, Issei Yoshida, Hiyori Yoshikawa, Koichiro Yoshino, Weiqiu You, Yang You, Zhiwen You, Bei Yu, Botao Yu, Changlong Yu, Charles Yu, Dezhi Yu, Dian Yu, Donglei Yu, Fanyang Yu, Fei Yu, Ge Yu, Guoxin Yu, Haiyang Yu, Hanchao Yu, Hang Yu, Hao Yu, Haofei Yu, Haoran Yu, Hong Yu, Hwanjo Yu, Jie Yu, Jun Yu, Junjie Yu, Lang Yu, Le Yu, Lu Yu, Nenghai Yu, Ping Yu, Qian Yu, Sangwon Yu, Seunguk Yu, Shi Yu, Shoubin Yu, Tan Yu, Tian Yu, Tianshu Yu, Tiezheng Yu, Weijiang Yu, Weijie Yu, Xiaobing Yu, Xiaodong Yu, Xiaoyan Yu, Yahan Yu, Yanchao Yu, Yangyang Yu, Zeping Yu, Zhang Yu, Zhi Yu, Zhongjing Yu, Zhuohao Yu, Caixia Yuan, Chenhan Yuan, Honglin Yuan, Jiahao Yuan,

Jiaqing Yuan, Jiayi Yuan, Jinjie Yuan, Michelle Yuan, Nicholas Jing Yuan, Peiwen Yuan, Shuzhou Yuan, Siyu Yuan, Weizhe Yuan, Xiangchi Yuan, Ye Yuan, Yige Yuan, Youliang Yuan, Yun-Hao Yuan, Zhangdie Yuan, Zheng Yuan, Zhihang Yuan, Ling Yue, Murong Yue, Shengbin Yue, Xian-gyu Yue, Yuanhao Yue, Zihao Yue, Hyeongu Yun, Hyokun Yun, JungMin Yun, Sangdoo Yun, Se-Young Yun, Seongjun Yun, Taewon Yun, Tian Yun, Lan Yuqian, François Yvon

Amir Zadeh, Matei Zaharia, Hamada Zahera, Iqra Zahid, Pardis Sadat Zahraei, Osmar Zaiane, Mo-hammadi Zaki, Kerem Zaman, Changtong Zan, Daoguang Zan, Hongying Zan, Yuhang Zang, An-drea Zaninello, Klim Zaporozets, Fadi A Zaraket, Alessandra Zarcone, Sina Zarrieß, Armel Randy Zebaze, Daniel Dajun Zeng, Daojian Zeng, Fengzhu Zeng, Hongchuan Zeng, Huimin Zeng, Jay-den Zeng, Jiali Zeng, Jichuan Zeng, Jingjie Zeng, Peijian Zeng, Peter Zeng, Qi Zeng, Qingcheng Zeng, Qingkai Zeng, Shenglai Zeng, Weixin Zeng, Zhichen Zeng, Zhiyuan Zeng, Ding Zepeng, Yang Zerui, Chrysoula Zerva, George Zerveas, Torsten Zesch, Daren Zha, ChengXiang Zhai, Fei-fei Zhai, Weihe Zhai, Zenan Zhai, Haolan Zhan, Huixin Zhan, Li-Ming Zhan, Mingjie Zhan, Pengwei Zhan, Qiusi Zhan, Runzhe Zhan, An Zhang, Binyan Zhang, Bo Zhang, Bowei Zhang, Bowen Zhang, Byoung-Tak Zhang, Caiqi Zhang, Chao Zhang, Chen Zhang, Chengruidong Zhang, Chenhao Zhang, Chenwei Zhang, Chenyang Zhang, Chi Zhang, Chiyu Zhang, Chiyuan Zhang, Chong Zhang, Chuanyi Zhang, Chunhui Zhang, Chunliang Zhang, Congzhi Zhang, Dan Zhang, Dell Zhang, Dong Zhang, Dongmei Zhang, Duzhen Zhang, Fan Zhang, Feng Zhang, Fengji Zhang, Fu Zhang, Fuwei Zhang, Fuzheng Zhang, Gaifan Zhang, Guchun Zhang, Haibo Zhang, Haichao Zhang, Hainan Zhang, Hanlei Zhang, Hanqing Zhang, Hao Zhang, Haoran Ranran Zhang, Haoyu Zhang, Hefan Zhang, Hongbin Zhang, Hongbo Zhang, Hongming Zhang, Hongyu Zhang, Hong-zhi Zhang, Huan Zhang, Huishuai Zhang, Huixuan Zhang, Ji Zhang, Jiajie Zhang, Jianfei Zhang, Jianguo Zhang, Jiani Zhang, Jianshu Zhang, Jiarui Zhang, Jie Zhang, Jihai Zhang, Jincheng Zhang, Jinchuan Zhang, Jingfeng Zhang, Jinghui Zhang, Jingyu Zhang, Jintian Zhang, Jipeng Zhang, Ji-tian Zhang, Jiwen Zhang, Jizhi Zhang, Jue Zhang, Juntian Zhang, Junyu Zhang, Junzhe Zhang, Kai Zhang, Kaike Zhang, Kaipeng Zhang, Kaiyan Zhang, Ke Zhang, Kun Zhang, Kunli Zhang, Lan Zhang, Le Zhang, Lefei Zhang, Lei Zhang, Liang Zhang, Licheng Zhang, Lihua Zhang, Lin Zhang, Lingxi Zhang, Lingling Zhang, Linhai Zhang, Linlin Zhang, Litian Zhang, Liwen Zhang, Longyin Zhang, Mengdi Zhang, Mengqi Zhang, Mi Zhang, Miaoran Zhang, Michael J Zhang, Mike Zhang, Min Zhang, Ming Zhang, Mozhi Zhang, Pei Zhang, Peinan Zhang, Peitian Zhang, Peng Zhang, Qi Zhang, Qiang Zhang, Qiming Zhang, Qin Zhang, Qing Zhang, Qinggang Zhang, Qingjie Zhang, Qinglin Zhang, Qingru Zhang, Qiyuan Zhang, Richong Zhang, Rongsheng Zhang, Rui Zhang, Ruike Zhang, Ruiyi Zhang, Ruochen Zhang, Ruqing Zhang, Sai Qian Zhang, Sheng Zhang, Shengyu Zhang, Shiyuan Zhang, Situo Zhang, Siyue Zhang, Songming Zhang, Songyang Zhang, Tianhua Zhang, Tianhui Zhang, Tianshu Zhang, Tianwei Zhang, Tianyi Zhang, Wei Zhang, Wei Emma Zhang, Weichen Zhang, Wen Zhang, Wenjun Zhang, Wentao Zhang, Wenyan Zhang, Wenzheng Zhang, Xi Zhang, Xianchao Zhang, Xiang Zhang, Xiangyu Zhang, Xiantao Zhang, Xiao Zhang, Xiaofan Zhang, Xiaofeng Zhang, Xiaojun Zhang, Xiaokun Zhang, Xiaotong Zhang, Xiaowang Zhang, Xin Zhang, Xinbo Zhang, Xinghua Zhang, Xinliang Frederick Zhang, Xinyang Zhang, Xinyu Zhang, Xinyuan Zhang, Xiuzhen Zhang, Xu Zhang, Xuanming Zhang, Xue Zhang, Xuejie Zhang, Xueyao Zhang, Xuhong Zhang, Xulang Zhang, Ya Zhang, Yan Zhang, Yancheng Zhang, Yang Zhang, Yanghai Zhang, Yangjun Zhang, Yanjian Zhang, Yanyong Zhang, Yanzhe Zhang, Yao Zhang, Yaping Zhang, Ye Zhang, Yi Zhang, Yice Zhang, Yichang Zhang, Yichao Zhang, Yichi Zhang, Yidan Zhang, Yifan Zhang, Yifei Zhang, Yihe Zhang, Yikai Zhang, Yiming Zhang, Yin Zhang, Ying Zhang, Yingji Zhang, Yingxue Zhang, Yingying Zhang, Yiqun Zhang, Yizhuo Zhang, Yongdong Zhang, Yongfeng Zhang, Yongheng Zhang, Yongqi Zhang, You Zhang, Yu Zhang, Yuansen Zhang, Yuanxing Zhang, Yuanzhe Zhang, Yue Zhang, Yuhan Zhang, Yuhao Zhang, Yuji Zhang, Yujie Zhang, Yunjian Zhang, Yunxiang Zhang, Yunyi Zhang, Yuqi Zhang, Yusen Zhang, Yuwei Zhang, Zecheng Zhang, Zekai Zhang, Zeliang Zhang, Zhao Zhang, Zhao-han Zhang, Zhehao Zhang, Zhen Zhang, Zhen-Ru Zhang, Zhexin Zhang, Zheyuan Zhang, Zhi

Zhang, Zhihan Zhang, Zhihao Zhang, Zhiqiang Zhang, Zhisong Zhang, Zhong Zhang, Zhongping Zhang, Zihan Zhang, Ziheng Zhang, Zihong Zhang, Zijian Zhang, Ziqi Zhang, Ziyin Zhang, Baining Zhao, Bowen Zhao, Fei Zhao, Feng Zhao, Guangxiang Zhao, Guangzhen Zhao, Guoshuai Zhao, Haiteng Zhao, Haiyan Zhao, Handong Zhao, Haozhe Zhao, Hui Zhao, Jiaxu Zhao, Jie Zhao, Jin Zhao, Jinman Zhao, Jinming Zhao, Junbo Zhao, Junchen Zhao, Kai Zhao, Kaiqi Zhao, Kaiyan Zhao, Kun Zhao, Liang Zhao, Libo Zhao, Lili Zhao, Lingchen Zhao, Lingjun Zhao, Mengyu Zhao, Minyi Zhao, Peilin Zhao, Pu Zhao, Qinghua Zhao, Qinlin Zhao, Ruihui Zhao, Ruochen Zhao, Sanqiang Zhao, Sendong Zhao, Shangqing Zhao, Shiwan Zhao, Shuai Zhao, Tiancheng Zhao, Weilin Zhao, Weixiang Zhao, Wenlong Zhao, Wenting Zhao, Xiangyu Zhao, Xiaoyan Zhao, Xingmeng Zhao, Xinpeng Zhao, Xinyu Zhao, Xiutian Zhao, Xujiang Zhao, Yang Zhao, Yang Yang Zhao, Yanyan Zhao, Yilun Zhao, Yiran Zhao, Yiren Zhao, Yu Zhao, Yufeng Zhao, Yukun Zhao, Yunxiao Zhao, Yusheng Zhao, Yutian Zhao, Yuze Zhao, Zhe Zhao, Zhendong Zhao, Zheng Zhao, Zhou Zhao, Zijian Zhao, Bi Zhen, Ranran Zhen, Boyuan Zheng, Chen Zheng, Chuan-nyang Zheng, Guangmin Zheng, Hai-Tao Zheng, Huanran Zheng, Jiyang Zheng, Jonathan Zheng, Kening Zheng, Mengxin Zheng, Mingqian Zheng, Mingyu Zheng, Qi Zheng, Rui Zheng, Tianshi Zheng, Tianyu Zheng, Tong Zheng, Weiguo Zheng, Xiaoqing Zheng, Yefeng Zheng, Yi Zheng, Yinhe Zheng, Yongsen Zheng, Yuanhang Zheng, Zhi Zheng, Zibin Zheng, Wen Zhiquan, Bai Zhi-xin, Guojin Zhong, Jialun Zhong, Jianyuan Zhong, Liangjun Zhong, Meizhi Zhong, Mian Zhong, Mingze Zhong, Qihuang Zhong, Shaochen Zhong, Weihong Zhong, Xian Zhong, Yang Zhong, Yiwu Zhong, Yujie Zhong, Aimin Zhou, Anda Zhou, Aojun Zhou, Bin Zhou, Changzhi Zhou, Cheng-Jie Zhou, Chuang Zhou, Dewei Zhou, Dongzhuoran Zhou, Enyu Zhou, Gengze Zhou, Hao Zhou, He Zhou, Houquan Zhou, Huachi Zhou, Huichi Zhou, Huixue Zhou, Jiayu Zhou, Jinfeng Zhou, Jingren Zhou, Jizhe Zhou, Jun Zhou, Junkai Zhou, Junsheng Zhou, Kaiji Zhou, Kun Zhou, Li Zhou, Mengyu Zhou, Mingyang Zhou, Mingyao Zhou, Naitian Zhou, Pan Zhou, Peilin Zhou, S Kevin Zhou, Shijia Zhou, Shilin Zhou, Shuchang Zhou, Shuigeng Zhou, Shuyan Zhou, Tong Zhou, Wei Zhou, Weixiao Zhou, Wengang Zhou, Xiabing Zhou, Xian Zhou, Xiangyang Zhou, Xiaobing Zhou, Xin Zhou, Xinyu Zhou, Xuhui Zhou, Yan Zhou, Yang Zhou, Yaqian Zhou, Yijie Zhou, Yin Zhou, Ying Zhou, Yiyang Zhou, Yiyun Zhou, Yongxin Zhou, Yu Zhou, Yue Zhou, Yufa Zhou, Yuhang Zhou, Yuhao Zhou, Yujun Zhou, Yukai Zhou, Yunhua Zhou, Yuqi Zhou, Zhanhui Zhou, Zhanke Zhou, Zhenhong Zhou, Zhong Zhou, Zihan Zhou, Andrew Zhu, Bin Benjamin Zhu, Chiwei Zhu, Chunjiang Zhu, Conghui Zhu, Dawei Zhu, Derui Zhu, Dongsheng Zhu, Fangqi Zhu, Hangxiao Zhu, He Zhu, Henghui Zhu, Hongyu Zhu, Huaisheng Zhu, Jia Zhu, Jian Zhu, Jianke Zhu, Jilong Zhu, Jing Zhu, JingBo Zhu, Jinhua Zhu, Junda Zhu, Kenny Q. Zhu, Kun Zhu, Kunlun Zhu, Linchao Zhu, Luyao Zhu, Mengna Zhu, Minfeng Zhu, Ming Zhu, Minjun Zhu, Peide Zhu, Qi Zhu, Qingfu Zhu, Qinglin Zhu, Rongzhi Zhu, Rui-Jie Zhu, Shaolin Zhu, Shengqi Zhu, Shucheng Zhu, Song-Chun Zhu, Suyang Zhu, Tinghui Zhu, Tong Zhu, Wang Zhu, Xiangru Zhu, Xiao Zhu, Xiaofei Zhu, Yangfu Zhu, Yaochen Zhu, Yi Zhu, Yingjie Zhu, Yonghua Zhu, Yutao Zhu, Yuxuan Zhu, Yuying Zhu, Zhihao Zhu, Zhihong Zhu, Ziwei Zhu, Zixiao Zhu, Haojie Zhuang, Haomin Zhuang, Huiping Zhuang, Shengyao Zhuang, Yuan Zhuang, Yuchen Zhuang, Yueting Zhuang, Zirui Zhuang, Ziyu Zhuang, Hai Zhuge, Jingming Zhuo, Caleb Ziems, Brian Douglas Zimmerman, Imed Zitouni, Xu Zixiang, Morteza Ziyadi, Stefan Zohren, Chang Zong, Linlin Zong, Qing Zong, Ruohan Zong, Bocheng Zou, Bowei Zou, Henry Peng Zou, Heqing Zou, Huiqi Zou, James Zou, Kai Zou, Lixin Zou, Yuexian Zou, Zhuoyang Zou, Vilém Zouhar, Guido Zuccon, Andrea Zugarini, Zayd Muhammad Kawakibi Zuhri, Liang Zujie, Pratibha Zunjare, Fei Zuo, Jialung Zuo, Max Zuo, Pierre Zweigenbaum, Hu zhang

## Outstanding Reviewers/Area Chairs

Nishant Balepur

Jonathan P. Chang, Jiaqi Chen, Yuxuan Chen, Daixuan Cheng, Liu Chengwu, Munmun De Choudhury

Suyash Damle, Tong Ding, Kevin Duh, Esra Dönmez

Alon Eirew

Giacomo Frisoni

Yeyun Gong

Chun-Ying Huang

Arkadiusz Janz, Zhenzhou Ji, Jiho Jin

Youna Kim

Matthieu Labeau, Xiwen Liang

Jiawei Ma, Mahan Malihi, Joel Ruben Antony Moniz, Sahal Shaji Mullappilly

Yaswanth Narsupalli, Long HB Nguyen, Jingcheng Niu

Tanmay Parekh, Dominic Petrak

Huachuan Qiu

Pretam Ray, Alla Rozovskaya

Andrey Sakhovskiy, Yves Scherrer, Vered Shwartz, Răzvan-Alexandru Smădu

Kavin R V, Justin Vasselli

Di Wang, Huimu Wang, Jiayi Wang, Mengru Wang, Ruida Wang

Tiankai Yang, Xiulin Yang, Zhicheng Yang, Wei Yao, Mun Yong Yi

Kerem Zaman, George Zerveas, Shuai Zhao, Wei Zhao, Shaolin Zhu, Qing Zong

# Keynote Talk

## No more Processing. Time to Discover.

Heng Ji

Siebel School of Computing and Data Science, University of Illinois



**Wednesday, November 5th – Time: from 09:30 to 10:30 – Room: B301**

**Abstract:** Scientific discovery, especially for new drugs and materials, urgently needs our help. The traditional manual approach is highly artisanal, and thus slow and expensive. Most importantly, many commercial drugs or materials have well-documented limitations that have remained unaddressed. In fact, AI for Science has become a rapidly growing field, with NLP researchers playing a central role by organizing shared tasks and competing on various leaderboards, especially through approaches powered by large language models (LLMs). In this talk, however, I will offer a provocative perspective on why our community may be heading in the wrong direction. Take drug discovery as an example: much of the existing work merely classifies properties of known drugs and thus discovers nothing, or generates molecules that are chemically impossible to make.

Unlike machines, human scientists are inherently “multilingual,” seamlessly navigating diverse modalities, from natural language in literature to complex scientific data such as molecular structures in knowledge bases. Human scientists also “think before they talk”: they ground their reasoning in deliberate reflection and subject new ideas to critical evaluation and verification. Against this backdrop, I argue that there is a correctable fundamental mismatch between the way LLMs work and the way scientists traditionally discover and verify new research hypotheses. I propose to design new LLM paradigms by drawing inspiration from the scientific discovery process itself: (1) “Observe” - acquire, represent and integrate knowledge from multiple data modalities; (2) “Think” - think critically to generate hypotheses; and (3) “Propose and Verify” - verify hypotheses through the Physical World.

As a prototype example, I will present mCLM, a modular Chemical-Language Model that speaks two complementary languages: one that represents molecular building blocks indicative of specific functions and compatible with automated modular assembly, and another that describes these functions in natural language. Experiments on 430 FDA-approved drugs showed that mCLM is capable of significantly improving chemical functions critical to determining drug potentials. mCLM, with only 3B parameters, also achieves improvements in function scores and synthetic accessibility relative to 7 other leading generative AI methods including GPT-5. mCLM can also reason on multiple functions and iteratively self-improve to rescue drug candidates that failed late in clinical trials (“fallen angels”). Preliminary animal testing results further underscore the promise of this approach. In the long term, I envision a comprehensive, multi-agent, human-in-the-loop autonomous laboratory, structured around iterative cycles of reasoning, proposal, synthesis, physical testing, feedback, and reasoning to enable never-ending self-improvement and co-evolution with human scientists.

**Bio:** Heng Ji is a Professor of Computer Science at Siebel School of Computing and Data Science, and a faculty member affiliated with Electrical and Computer Engineering Department, Coordinated Science Laboratory, and Carl R. Woese Institute for Genomic Biology of University of Illinois Urbana-Champaign. She is an Amazon Scholar. She is the Founding Director of Amazon-Illinois Center on AI for Interactive Conversational Experiences (AICE), and the Founding Director of CapitalOne-Illinois Center on AI Safety and Knowledge Systems (ASKS). She received Ph.D. in Computer Science from New York University. Her research interests focus on Natural Language Processing, especially on Multimedia Multilingual Information Extraction, Knowledge-enhanced Large Language Models and Vision-Language Models, AI for Science, and Science-inspired AI. The awards she received include Outstanding Paper Award at ACL2024, two Outstanding Paper Awards at NAACL2024, “Young Scientist” by the World Laureates Association in 2023 and 2024, “Young Scientist” and a member of the Global Future Council on the Future of Computing by the World Economic Forum in 2016 and 2017, “Women Leaders of Conversational AI” (Class of 2023) by Project Voice, “AI’s 10 to Watch” Award by IEEE Intelligent Systems in 2013, NSF CAREER award in 2009, PACLIC2012 Best paper runner-up, “Best of ICDM2013” paper award, “Best of SDM2013” paper award, ACL2018 Best Demo paper nomination, ACL2020 Best Demo Paper Award, NAACL2021 Best Demo Paper Award, Google Research Award in 2009 and 2014, IBM Watson Faculty Award in 2012 and 2014 and Bosch Research Award in 2014-2018. She has coordinated the NIST TAC Knowledge Base Population task 2010-2020. She served as the associate editor for IEEE/ACM Transaction on Audio, Speech, and Language Processing, and the Program Committee Co-Chair of many conferences including NAACL-HLT2018 and AACL-IJCNLP2022. She was elected as the North American Chapter of the Association for Computational Linguistics (NAACL) secretary 2020-2023.

**Keynote Talk**  
**Looking at NLP through the lens of the (computational)  
social sciences**

**Jana Diesner**  
Technical University of Munich



**Thursday, November 6th – Time: from 09:00 to 10:00 – Room: B301**

**Abstract:** Using empirical methods from NLP has led to tremendous progress in the social sciences and humanities. I address how leveraging innovative NLP techniques has helped to discover and validate hypotheses about the functioning and patterns of complex societal systems, jointly considering the content and structure of social and socio-technical interactions, and scaling up the analysis of human-centered, qualitative, and multi-modal data. Knowledge from the social sciences has in turn informed and advanced data annotation, model building, and result explanation in NLP. At the same time, the choices we make when studying and predicting the structure and behavior of groups matter. I discuss how differences in research design, epistemology, and culturally contextualized algorithms impact knowledge discovery, theory development, practical applications, and policies. I conclude with implications of work at the nexus of NLP and the social sciences on innovation, validation, and bridging differences in scales.

**Bio:** Jana Diesner is a Full Professor at the Technical University of Munich, School of Social Science and Technology, with a joint appointment at the School of Computation, Information and Technology. She leads the Human-Centered Computing group. Her group works on methods from network science, natural language processing, machine learning, and AI, and integrates them with theories from the social sciences and humanities to advance our understanding of theories and patterns of complex societal systems. Their work also considers cultural contexts and ethical concerns to advance responsibility in computational sciences. Before joining TUM in 2024, Jana was a tenured professor at the University of Illinois Urbana Champaign, School of Information Sciences. She earned her Ph.D. at Carnegie Mellon, School of Computer Science.



# Keynote Talk

## Open-Science AI: Building Language, Vision, and Reasoning Models that Drive Innovation

Hannaneh Hajishirzi  
University of Washington



**Friday, November 7th – Time: from 09:00 to 10:00 – Room: B301**

**Abstract:** Innovation and competitiveness in AI are driven not only by building powerful systems but also by creating enduring, extensible solutions that empower the research community to explore, evaluate, and build upon them further. In this talk, I discuss how to develop open language, vision, and reasoning models that can compete with proprietary counterparts—models that are fully open and scientifically grounded. I will highlight how these open-science AI efforts have led to innovative strategies for data curation, new paradigms in model training, improved reasoning algorithms, and advances in model traceability, among others. These ideas will be illustrated through examples from my team’s work on OLMo, Tulu, Molmo, FlexOLMo, and OLMoTrace, as well as related contributions from the broader research community.

More specifically, I will share lessons learned from building our fully open language models, OLMo, demonstrating that with the right scientific and engineering approaches, it is possible to refine model capabilities and achieve state-of-the-art performance. I will explain how this foundation has enabled us to develop FlexOLMo, a new paradigm for modular language model training that supports both data privacy and data collaboration, and OLMoTrace, a tool for tracing model outputs back to their training data. I will also discuss RLVR, our new approach to reinforcement learning with verifiable rewards in Tulu 3, and how it has led to the creation of stronger reasoning models, catalyzing an active new area of research in verifiable reasoning. Finally, I will highlight Dolma and PixMo, illustrating how novel data curation and mixing strategies further enhance model capabilities and enable effective, high-quality training of language and vision models.

**Bio:** Hanna Hajishirzi is a Professor of Computer Science at the University of Washington and a Senior Director of AI at AI2. Her research spans generative AI and natural language processing, with a focus on building pioneering, open-science AI solutions. She co-leads the OLMo and Tulu projects, advancing fully open language and reasoning models to accelerate the science of AI, empower the research community, and champion openness as a driver of innovation. These models have been downloaded more than 10 million times as of 2025 and were recognized with GeekWire’s Innovation of the Year award. She is a co-PI of a \$152M NSF- and NVIDIA-supported grant to develop the next generation of open multimodal models.

She is a recipient of the Sloan Fellowship (2021), the Uncommon Thinker Award (2025), the NSF CAREER Award (2021), Torode family Career development professorship (2022), the Allen Distinguished

Investigator Award (2014), the UIUC Alumni Award (2024), and was a finalist for the VentureBeat Women in AI Award (2024, 2025). Her research has earned recognition at leading venues, with papers receiving or being finalists for awards at ACL 2025, CVPR 2025, ACL 2024 (Best Paper and Best Resource Paper), CVPR 2022, AKBC 2020, and SIGDIAL 2012.

# Table of Contents

|                                                                                                                                                                                                                          |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Towards Automated Error Discovery: A Study in Conversational AI</i><br>Dominic Petrak, Thy Thy Tran and Iryna Gurevych .....                                                                                          | 1   |
| <i>Break the Checkbox: Challenging Closed-Style Evaluations of Cultural Alignment in LLMs</i><br>Mohsinul Kabir, Ajwad Abrar and Sophia Ananiadou .....                                                                  | 25  |
| <i>Biased Tales: Cultural and Topic Bias in Generating Children's Stories</i><br>Donya Rooein, Vilém Zouhar, Debora Nozza and Dirk Hovy .....                                                                            | 53  |
| <i>Large Language Models as Realistic Microservice Trace Generators</i><br>Donghyun Kim, Sriram Ravula, Taemin Ha, Alex Dimakis, Daehyeok Kim and Aditya Akella                                                          | 74  |
| <i>JUDGEBERT: Assessing Legal Meaning Preservation Between Sentences</i><br>David Beauchemin, Michelle Albert-Rochette, Richard Khoury and Pierre-Luc Déziel .....                                                       | 93  |
| <i>QFrCoLA: a Quebec-French Corpus of Linguistic Acceptability Judgments</i><br>David Beauchemin and Richard Khoury .....                                                                                                | 120 |
| <i>Revisiting LLM Value Probing Strategies: Are They Robust and Expressive?</i><br>Siqi Shen, Mehar Singh, Lajanugen Logeswaran, Moontae Lee, Honglak Lee and Rada Mihalcea                                              | 132 |
| <i>A Systematic Analysis of Base Model Choice for Reward Modeling</i><br>Kian Ahrabian, Pegah Jandaghi, Negar Mokhberian, Sai Praneeth Karimireddy and Jay Pujara                                                        | 147 |
| <i>Comparing Specialised Small and General Large Language Models on Text Classification: 100 Labeled Samples to Achieve Break-Even Performance</i><br>Branislav Pecher, Ivan Srba and Maria Bielikova .....              | 166 |
| <i>Is the Top Still Spinning? Evaluating Subjectivity in Narrative Understanding</i><br>Melanie Subbiah, Akankshya Mishra, Grace Kim, Liyan Tang, Greg Durrett and Kathleen McKeown .....                                | 186 |
| <i>MathTutorBench: A Benchmark for Measuring Open-ended Pedagogical Capabilities of LLM Tutors</i><br>Jakub Macina, Nico Daheim, Ido Hakimi, Manu Kapur, Iryna Gurevych and Mrinmaya Sachan                              | 205 |
| <i>Preemptive Detection and Correction of Misaligned Actions in LLM Agents</i><br>Haishuo Fang, Xiaodan Zhu and Iryna Gurevych .....                                                                                     | 223 |
| <i>Fingerprinting LLMs through Survey Item Factor Correlation: A Case Study on Humor Style Questionnaire</i><br>Simon Munker .....                                                                                       | 246 |
| <i>Gradient-Attention Guided Dual-Masking Synergetic Framework for Robust Text-based Person Retrieval</i><br>Tianlu Zheng, Yifan Zhang, Xiang An, Ziyong Feng, Kaicheng Yang and Qichuan Ding ....                       | 260 |
| <i>From Problem-Solving to Teaching Problem-Solving: Aligning LLMs with Pedagogy using Reinforcement Learning</i><br>David Dinucu-Jianu, Jakub Macina, Nico Daheim, Ido Hakimi, Iryna Gurevych and Mrinmaya Sachan ..... | 273 |

|                                                                                                                                 |     |
|---------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>CompKBQA: Component-wise Task Decomposition for Knowledge Base Question Answering</i>                                        |     |
| Yuhang Tian, Dandan Song, Zhijing Wu, Pan Yang, Changzhi Zhou, Jun Yang, Hao Wang, Hui-peng Ma, Chenhao Li and Luan Zhang ..... | 294 |
| <i>Permutative Preference Alignment from Listwise Ranking of Human Judgments</i>                                                |     |
| Yang Zhao, Yixin Wang and Mingzhang Yin.....                                                                                    | 311 |
| <i>ToneCraft: Cantonese Lyrics Generation with Harmony of Tones and Pitches</i>                                                 |     |
| Junyu Cheng, Chang Pan and Shuangyin Li.....                                                                                    | 336 |
| <i>SensorLLM: Aligning Large Language Models with Motion Sensors for Human Activity Recognition</i>                             |     |
| Zeichen Li, Shohreh Deldari, Linyao Chen, Hao Xue and Flora D. Salim .....                                                      | 355 |
| <i>MixLoRA-DSI: Dynamically Expandable Mixture-of-LoRA Experts for Rehearsal-Free Generative Retrieval over Dynamic Corpora</i> |     |
| Tuan-Luc Huynh, Thuy-Trang Vu, Weiqing Wang, Trung Le, Dragan Gasevic, Yuan-Fang Li and Thanh-Toan Do .....                     | 381 |
| <i>ViClaim: A Multilingual Multilabel Dataset for Automatic Claim Detection in Videos</i>                                       |     |
| Patrick Giedemann, Pius von Däniken, Jan Milan Deriu, Alvaro Rodrigo, Anselmo Peñas and Mark Cieliebak .....                    | 398 |
| <i>DeepResearcher: Scaling Deep Research via Reinforcement Learning in Real-world Environments</i>                              |     |
| Yuxiang Zheng, Dayuan Fu, Xiangkun Hu, Xiaojie Cai, Lyumanshan Ye, Pengrui Lu and Pengfei Liu .....                             | 415 |
| <i>Mixture of Length and Pruning Experts for Knowledge Graphs Reasoning</i>                                                     |     |
| Enjun Du, Siyi Liu and Yongqi Zhang.....                                                                                        | 433 |
| <i>MPRF: Interpretable Stance Detection through Multi-Path Reasoning Framework</i>                                              |     |
| ZhaoDan Zhang, Jin Zhang, Hui Xu, Jiafeng Guo and Xueqi Cheng .....                                                             | 455 |
| <i>Analyzing the Effects of Supervised Fine-Tuning on Model Knowledge from Token and Parameter Levels</i>                       |     |
| Junjie Ye, Yuming Yang, Yang Nan, Shuo Li, Qi Zhang, Tao Gui, Xuanjing Huang, Peng Wang, Zhongchao Shi and Jianping Fan.....    | 472 |
| <i>JI<sup>2</sup>S: Joint Influence-Aware Instruction Data Selection for Efficient Fine-Tuning</i>                              |     |
| Jingyu Wei, Bo Liu, Tianjiao Wan, Baoyun Peng, Xingkong Ma and Mengmeng Guo .....                                               | 515 |
| <i>SoundMind: RL-Incentivized Logic Reasoning for Audio-Language Models</i>                                                     |     |
| Xingjian Diao, Chunhui Zhang, Keyi Kong, Weiyi Wu, Chiyu Ma, Zhongyu Ouyang, Peijun Qing, Soroush Vosoughi and Jiang Gui.....   | 529 |
| <i>Seeing More, Saying More: Lightweight Language Experts are Dynamic Video Token Compressors</i>                               |     |
| Xiangchen Wang, Jinrui Zhang, Teng Wang, Haigang Zhang and Feng Zheng .....                                                     | 542 |
| <i>RoT: Enhancing Table Reasoning with Iterative Row-Wise Traversals</i>                                                        |     |
| Xuanliang Zhang, Dingzirui Wang, Keyan Xu, Qingfu Zhu and Wanxiang Che .....                                                    | 560 |
| <i>T-MAD: Target-driven Multimodal Alignment for Stance Detection</i>                                                           |     |
| ZhaoDan Zhang, Jin Zhang, Xueqi Cheng and Hui Xu .....                                                                          | 581 |
| <i>Emotion Transfer with Enhanced Prototype for Unseen Emotion Recognition in Conversation</i>                                  |     |
| Kun Peng, Cong Cao, Hao Peng, Guanlin Wu, Zhifeng Hao, Lei Jiang, Yanbing Liu and Philip S. Yu.....                             | 597 |

|                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>PBI-Attack: Prior-Guided Bimodal Interactive Black-Box Jailbreak Attack for Toxicity Maximization</i>                                |     |
| Ruoxi Cheng, Yizhong Ding, Shuirong Cao, Ranjie Duan, Xiaoshuang Jia, Shaowei Yuan, Simeng Qin, Zhiqiang Wang and Xiaojun Jia .....     | 610 |
| <i>Training a Utility-based Retriever Through Shared Context Attribution for Retrieval-Augmented Language Models</i>                    |     |
| Yilong Xu, Jinhua Gao, Xiaoming Yu, Yuanhai Xue, Baolong Bi, Huawei Shen and Xueqi Cheng                                                | 630 |
| <i>SportReason: Evaluating Retrieval-Augmented Reasoning across Tables and Text for Sports Question Answering</i>                       |     |
| Kaiyue Feng, Siyue Zhang, Bingsen Chen, Yilun Zhao and Chen Zhao .....                                                                  | 650 |
| <i>MAC-Tuning: LLM Multi-Compositional Problem Reasoning with Enhanced Knowledge Boundary Awareness</i>                                 |     |
| Junsheng Huang, Zhitao He, Yuchen Huang, Sandeep Polisetty, Qingyun Wang and Yi R. Fung                                                 | 664 |
| <i>CODI: Compressing Chain-of-Thought into Continuous Space via Self-Distillation</i>                                                   |     |
| Zhenyi Shen, Hanqi Yan, Linhai Zhang, Zhanghao Hu, Yali Du and Yulan He .....                                                           | 678 |
| <i>PAFT: Prompt-Agnostic Fine-Tuning</i>                                                                                                |     |
| Chenxing Wei, Yao Shu, Mingwen Ou, Ying He and Fei Yu .....                                                                             | 695 |
| <i>Theorem-Validated Reverse Chain-of-Thought Problem Generation for Geometric Reasoning</i>                                            |     |
| Deng Linger, Linghao Zhu, Yuliang Liu, Yu Wang, Qunyi Xie, Jingjing Wu, Gang Zhang, Yingying Zhu and Xiang Bai .....                    | 719 |
| <i>TACO: Enhancing Multimodal In-context Learning via Task Mapping-Guided Sequence Configuration</i>                                    |     |
| Yanshu Li, Jianjiang Yang, Tian Yun, Pinyuan Feng, Jinfa Huang and Ruixiang Tang .....                                                  | 737 |
| <i>Towards Controllable Speech Synthesis in the Era of Large Language Models: A Systematic Survey</i>                                   |     |
| Tianxin Xie, Yan Rong, Pengfei Zhang, Wenwu Wang and Li Liu .....                                                                       | 765 |
| <i>Automating Steering for Safe Multimodal Large Language Models</i>                                                                    |     |
| Lyucheng Wu, Mengru Wang, Ziwen Xu, Tri Cao, Nay Oo, Bryan Hooi and Shumin Deng ..                                                      | 793 |
| <i>EMNLP: Educator-role Moral and Normative Large Language Models Profiling</i>                                                         |     |
| Yilin Jiang, Mingzi Zhang, Sheng Jin, Zengyi Yu, Xiangjie Kong and Binghao Tu .....                                                     | 816 |
| <i>TracSum: A New Benchmark for Aspect-Based Summarization with Sentence-Level Traceability in Medical Domain</i>                       |     |
| Bohao Chu, Meijie Li, Sameh Frihat, Chengyu Gu, Georg Lodde, Elisabeth Livingstone and Norbert Fuhr .....                               | 845 |
| <i>Context Reasoner: Incentivizing Reasoning Capability for Contextualized Privacy and Safety Compliance via Reinforcement Learning</i> |     |
| Wenbin Hu, Haoran Li, Huihao Jing, Qi Hu, Ziqian Zeng, Sirui Han, Xu Heli, Tianshu Chu, Peizhao Hu and Yangqiu Song .....               | 866 |
| <i>Towards General-Domain Word Sense Disambiguation: Distilling Large Language Model into Compact Disambiguator</i>                     |     |
| Liqiang Ming, Sheng-hua Zhong and Yuncong Li .....                                                                                      | 885 |
| <i>SLoW: Select Low-frequency Words! Automatic Dictionary Selection for Translation on Large Language Models</i>                        |     |
| Hongyuan Lu, Zixuan Li, Zefan Zhang and Wai Lam .....                                                                                   | 899 |

|                                                                                                                                                          |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Parallel Continuous Chain-of-Thought with Jacobi Iteration</i>                                                                                        |      |
| Haoyi Wu, Zhihao Teng and Kewei Tu .....                                                                                                                 | 915  |
| <i>EQA-RM: A Generative Embodied Reward Model with Test-time Scaling</i>                                                                                 |      |
| Yuhang Chen, Zhen Tan and Tianlong Chen .....                                                                                                            | 928  |
| <i>Refusal-Aware Red Teaming: Exposing Inconsistency in Safety Evaluations</i>                                                                           |      |
| Yongkang Chen, Xiaohu Du, Xiaotian Zou, Chongyang Zhao, Huan Deng, Hu LI and Xiaohui Kuang .....                                                         | 947  |
| <i>OmniThink: Expanding Knowledge Boundaries in Machine Writing through Thinking</i>                                                                     |      |
| Zekun Xi, Wenbiao Yin, Jizhan Fang, Jialong Wu, Runnan Fang, Yong Jiang, Pengjun Xie, Fei Huang, Huajun Chen and Ningyu Zhang .....                      | 957  |
| <i>LinkAlign: Scalable Schema Linking for Real-World Large-Scale Multi-Database Text-to-SQL</i>                                                          |      |
| Yihan Wang, Peiyu Liu and Xin Yang .....                                                                                                                 | 978  |
| <i>On Relation-Specific Neurons in Large Language Models</i>                                                                                             |      |
| Yihong Liu, Runsheng Chen, Lea Hirlimann, Ahmad Dawar Hakimi, Mingyang Wang, Amir Hossein Kargar, Sascha Rothe, François Yvon and Hinrich Schuetze ..... | 993  |
| <i>IPIGuard: A Novel Tool Dependency Graph-Based Defense Against Indirect Prompt Injection in LLM Agents</i>                                             |      |
| Hengyu An, Jinghui Zhang, Tianyu Du, Chunyi Zhou, Qingming Li, Tao Lin and Shouling Ji                                                                   | 1024 |
| <i>ProtoVQA: An Adaptable Prototypical Framework for Explainable Fine-Grained Visual Question Answering</i>                                              |      |
| Xingjian Diao, Weiyi Wu, Keyi Kong, Peijun Qing, Xinwen Xu, Ming Cheng, Soroush Vosoughi and Jiang Gui .....                                             | 1041 |
| <i>SEA: Supervised Embedding Alignment for Token-Level Visual-Textual Integration in MLLMs</i>                                                           |      |
| Yuanyang Yin, Yaqi Zhao, Yajie Zhang, Yuanxing Zhang, Ke Lin, Jiahao Wang, Xin Tao, Pengfei Wan, Wentao Zhang and Feng Zhao .....                        | 1059 |
| <i>Molecular String Representation Preferences in Pretrained LLMs: A Comparative Study in Zero- &amp; Few-Shot Molecular Property Prediction</i>         |      |
| George Arthur Baker, Mario Sanz-Guerrero and Katharina von der Wense .....                                                                               | 1072 |
| <i>Weight-Aware Activation Sparsity with Constrained Bayesian Optimization Scheduling for Large Language Models</i>                                      |      |
| Ming Wang, Miao Zhang, Xuebo Liu and Liqiang Nie .....                                                                                                   | 1087 |
| <i>DatawiseAgent: A Notebook-Centric LLM Agent Framework for Adaptive and Robust Data Science Automation</i>                                             |      |
| Ziming You, Yumiao Zhang, Dexuan Xu, Yiwei Lou, Yandong Yan, Wei Wang, Huamin Zhang and Yu Huang .....                                                   | 1100 |
| <i>VC4VG: Optimizing Video Captions for Text-to-Video Generation</i>                                                                                     |      |
| Yang Du, Zhuoran Lin, Kaiqiang Song, Biao Wang, Zhicheng Zheng, Tiezheng Ge, Bo Zheng and Qin Jin .....                                                  | 1125 |
| <i>LaMP-QA: A Benchmark for Personalized Long-form Question Answering</i>                                                                                |      |
| Alireza Salemi and Hamed Zamani .....                                                                                                                    | 1140 |
| <i>The LLM Already Knows: Estimating LLM-Perceived Question Difficulty via Hidden Representations</i>                                                    |      |
| Yubo Zhu, Dongrui Liu, Zecheng Lin, Wei Tong, Sheng Zhong and Jing Shao .....                                                                            | 1161 |

|                                                                                                                                  |      |
|----------------------------------------------------------------------------------------------------------------------------------|------|
| <i>MCIP: Protecting MCP Safety via Model Contextual Integrity Protocol</i>                                                       |      |
| Huihao Jing, Haoran Li, Wenbin Hu, Qi Hu, Xu Heli, Tianshu Chu, Peizhao Hu and Yangqiu Song                                      | 1178 |
| <i>SAKI-RAG: Mitigating Context Fragmentation in Long-Document RAG via Sentence-level Attention Knowledge Integration</i>        |      |
| Wenyu Tao, Xiaofen Xing, Zeliang Li and Xiangmin Xu                                                                              | 1196 |
| <i>Skeletons Matter: Dynamic Data Augmentation for Text-to-Query</i>                                                             |      |
| Yuchen Ji, Bo Xu, Jie Shi, Jiaqing Liang, Deqing Yang, Yu Mao, Hai Chen and Yanghua Xiao                                         | 1215 |
| <i>CondenseLM: LLMs-driven Text Dataset Condensation via Reward Matching</i>                                                     |      |
| Cheng Shen, Yew-Soon Ong and Joey Tianyi Zhou                                                                                    | 1238 |
| <i>MovieCORE: COgnitive REasoning in Movies</i>                                                                                  |      |
| Gueter Josmy Faure, Min-Hung Chen, Jia-Fong Yeh, Ying Cheng, Hung-Ting Su, Yung-Hao Tang, Shang-Hong Lai and Winston H. Hsu      | 1254 |
| <i>Think Wider, Detect Sharper: Reinforced Reference Coverage for Document-Level Self-Contradiction Detection</i>                |      |
| Yuhao Chen, Yuanjie Lyu, Shuochen Liu, Chao Zhang, Junhui Lv and Tong Xu                                                         | 1274 |
| <i>DRISHTIKON: A Multimodal Multilingual Benchmark for Testing Language Models' Understanding on Indian Culture</i>              |      |
| Arijit Maji, Raghvendra Kumar, Akash Ghosh, Anushka, Nemil Shah, Abhilekh Borah, Vanshika Shah, Nishant Mishra and Sriparna Saha | 1290 |
| <i>LingGym: How Far Are LLMs from Thinking Like Field Linguists?</i>                                                             |      |
| Changbing Yang, Franklin Ma, Freda Shi and Jian Zhu                                                                              | 1315 |
| <i>Learning from Few Samples: A Novel Approach for High-Quality Malcode Generation</i>                                           |      |
| Haijian Ma, Daizong Liu, Xiaowen Cai, Pan Zhou and Yulai Xie                                                                     | 1342 |
| <i>Personality Matters: User Traits Predict LLM Preferences in Multi-Turn Collaborative Tasks</i>                                |      |
| Sarfaroze Yunusov, Kaige Chen, Kazi Nishat Anwar and Ali Emami                                                                   | 1360 |
| <i>VisualWebInstruct: Scaling up Multimodal Instruction Data through Web Search</i>                                              |      |
| Yiming Jia, Jiachen Li, Xiang Yue, Bo Li, Ping Nie, Kai Zou and Wenhua Chen                                                      | 1374 |
| <i>Thinking Out Loud: Do Reasoning Models Know When They're Right?</i>                                                           |      |
| Qingcheng Zeng, Weihao Xuan, Leyang Cui and Rob Voigt                                                                            | 1395 |
| <i>Seeing is Believing, but How Much? A Comprehensive Analysis of Verbalized Calibration in Vision-Language Models</i>           |      |
| Weihao Xuan, Qingcheng Zeng, Heli Qi, Junjue Wang and Naoto Yokoya                                                               | 1409 |
| <i>Enhancing Efficiency and Exploration in Reinforcement Learning for LLMs</i>                                                   |      |
| Mengqi Liao, Xiangyu Xi, Chen Ruinian, Jia Leng, Yangen Hu, Ke Zeng, Shuai Liu and Huaiyu Wan                                    | 1452 |
| <i>LLM Bias Detection and Mitigation through the Lens of Desired Distributions</i>                                               |      |
| Ingroj Shrestha and Padmini Srinivasan                                                                                           | 1465 |
| <i>MEBench: Benchmarking Large Language Models for Cross-Document Multi-Entity Question Answering</i>                            |      |
| Teng Lin, Yuyu Luo, Honglin Zhang, Jicheng Zhang, Chunlin Liu, Kaishun Wu and Nan Tang                                           | 1482 |



*POSITION BIAS MITIGATES POSITION BIAS: Mitigate Position Bias Through Inter-Position Knowledge Distillation*

Yifei Wang, Feng Xiong, Yong Wang, Linjing Li, Xiangxiang Chu and Daniel Dajun Zeng . 1496

*MMLU-ProX: A Multilingual Benchmark for Advanced Large Language Model Evaluation*

Weihao Xuan, Rui Yang, Heli Qi, Qingcheng Zeng, Yunze Xiao, Aosong Feng, Dairui Liu, Yun Xing, Junjue Wang, Fan Gao, Jinghui Lu, Yuang Jiang, Huitao Li, Xin Li, Kunyu Yu, Ruihai Dong, Shangding Gu, Yuekang Li, Xiaofei Xie, Felix Juefei-Xu, Foutse Khomh, Osamu Yoshie, Qingyu Chen, Douglas Teodoro, Nan Liu, Randy Goebel, Lei Ma, Edison Marrese-Taylor, Shijian Lu, Yusuke Iwasawa, Yutaka Matsuo and Irene Li ..... 1514

*NL-Debugging: Exploiting Natural Language as an Intermediate Representation for Code Debugging*

Weiming Zhang, Qingyao Li, Xinyi Dai, Jizheng Chen, Kounianhua Du, Weiwen Liu, Yasheng Wang, Ruiming Tang, Yong Yu and Weinan Zhang ..... 1534

*Persuasion Dynamics in LLMs: Investigating Robustness and Adaptability in Knowledge and Safety with DuET-PD*

Bryan Chen Zhengyu Tan, Daniel Wai Kit Chin, Zhengyuan Liu, Nancy F. Chen and Roy Ka-Wei Lee ..... 1551

*POINTS-Reader: Distillation-Free Adaptation of Vision-Language Models for Document Conversion*

Yuan Liu, Zhongyin Zhao, Le Tian, Haicheng Wang, Xubing Ye, Yangxiu You, Zilin Yu, Chuhan Wu, Zhou Xiao, Yang Yu and Jie Zhou ..... 1577

*Large Language Models for Automated Literature Review: An Evaluation of Reference Generation, Abstract Writing, and Review Composition*

Xuemei Tang, Xufeng Duan and Zhenguang Cai ..... 1603

*CoBia: Constructed Conversations Can Trigger Otherwise Concealed Societal Biases in LLMs*

Nafiseh Nikeghbal, Amir Hossein Kargaran and Jana Diesner ..... 1619

*From Schema to State: Zero-Shot Scheme-Only Dialogue State Tracking via Diverse Synthetic Dialogue and Step-by-Step Distillation*

Huan Xu, Zequn Li, Wen Tang and Jian Jun Zhang ..... 1641

*Beyond the Surface: Measuring Self-Preference in LLM Judgments*

Zhi-Yuan Chen, Hao Wang, Xinyu Zhang, Enrui Hu and Yankai Lin ..... 1654

*Beyond Input Activations: Identifying Influential Latents by Gradient Sparse Autoencoders*

Dong Shu, Xuansheng Wu, Haiyan Zhao, Mengnan Du and Ninghao Liu ..... 1674

*Utility-Focused LLM Annotation for Retrieval and Retrieval-Augmented Generation*

Hengran Zhang, Minghao Tang, Keping Bi, Jiafeng Guo, Shihao Liu, Daiting Shi, Dawei Yin and Xueqi Cheng ..... 1684

*CiteBART: Learning to Generate Citations for Local Citation Recommendation*

Ege Yiğit Çelik and Selma Tekir ..... 1704

*Autoformalization in the Wild: Assessing LLMs on Real-World Mathematical Definitions*

Lan Zhang, Marco Valentino and Andre Freitas ..... 1721

*Culture Cartography: Mapping the Landscape of Cultural Knowledge*

Caleb Ziems, William Barr Held, Jane Yu, Amir Goldberg, David Grusky and Diyi Yang .. 1740

*Interpretability Analysis of Arithmetic In-Context Learning in Large Language Models*

Gregory Polyakov, Christian Hepting, Carsten Eickhoff and Seyed Ali Bahrainian ..... 1759



|                                                                                                                                                          |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>SwarmAgentic: Towards Fully Automated Agentic System Generation via Swarm Intelligence</i>                                                            |      |
| Yao Zhang, Chenyang Lin, Shijie Tang, Haokun Chen, Shijie Zhou, Yunpu Ma and Volker Tresp                                                                |      |
|                                                                                                                                                          | 1779 |
| <i>We Politely Insist: Your LLM Must Learn the Persian Art of Taarof</i>                                                                                 |      |
| Nikta Gohari Sadr, Sahar Heidariasl, Karine Megerdooimian, Laleh Seyyed-Kalantari and Ali Emami                                                          |      |
|                                                                                                                                                          | 1820 |
| <i>Unstructured Evidence Attribution for Long Context Query Focused Summarization</i>                                                                    |      |
| Dustin Wright, Zain Muhammad Mujahid, Lu Wang, Isabelle Augenstein and David Jurgens                                                                     |      |
|                                                                                                                                                          | 1840 |
| <i>RAVEN: Query-Guided Representation Alignment for Question Answering over Audio, Video, Embedded Sensors, and Natural Language</i>                     |      |
| Subrata Biswas, Mohammad Nur Hossain Khan and Bashima Islam                                                                                              |      |
|                                                                                                                                                          | 1869 |
| <i>Cache-of-Thought: Master-Apprentice Framework for Cost-Effective Vision Language Model Reasoning</i>                                                  |      |
| Mingyuan Wu, Jize Jiang, Haozhen Zheng, Meitang Li, Zhaocheng Li, Beitong Tian, Bo Chen, Yongjoo Park, Minjia Zhang, ChengXiang Zhai and Klara Nahrstedt |      |
|                                                                                                                                                          | 1896 |
| <i>Video Compression Commander: Plug-and-Play Inference Acceleration for Video Large Language Models</i>                                                 |      |
| Xuyang Liu, Yiyu Wang, Junpeng Ma and Linfeng Zhang                                                                                                      |      |
|                                                                                                                                                          | 1911 |
| <i>Router-Tuning: A Simple and Effective Approach for Dynamic Depth</i>                                                                                  |      |
| Shwai He, Tao Ge, Guoheng Sun, Bowei Tian, Xiaoyang Wang and Dong Yu                                                                                     |      |
|                                                                                                                                                          | 1926 |
| <i>Foot-In-The-Door: A Multi-turn Jailbreak for LLMs</i>                                                                                                 |      |
| Zixuan Weng, Xiaolong Jin, Jinyuan Jia and Xiangyu Zhang                                                                                                 |      |
|                                                                                                                                                          | 1940 |
| <i>TurnaboutLLM: A Deductive Reasoning Benchmark from Detective Games</i>                                                                                |      |
| Yuan Yuan, Muyu He, Muhammad Adil Shahid, Ziyang Li, Jiani Huang and Li Zhang                                                                            |      |
|                                                                                                                                                          | 1952 |
| <i>Transferable Direct Prompt Injection via Activation-Guided MCMC Sampling</i>                                                                          |      |
| Minghui Li, Hao Zhang, Yechao Zhang, Wei Wan, Shengshan Hu, Pei Xiaobing and Jing Wang                                                                   |      |
|                                                                                                                                                          | 1967 |
| <i>Direct Judgement Preference Optimization</i>                                                                                                          |      |
| PeiFeng Wang, Austin Xu, Yilun Zhou, Caiming Xiong and Shafiq Joty                                                                                       |      |
|                                                                                                                                                          | 1980 |
| <i>WebInject: Prompt Injection Attack to Web Agents</i>                                                                                                  |      |
| Xilong Wang, John Bloch, Zedian Shao, Yuepeng Hu, Shuyan Zhou and Neil Zhenqiang Gong                                                                    |      |
|                                                                                                                                                          | 2011 |
| <i>F<sup>2</sup>Bench: An Open-ended Fairness Evaluation Benchmark for LLMs with Factuality Considerations</i>                                           |      |
| Tian Lan, Jiang Li, Yemin Wang, Xu Liu, Xiangdong Su and Guanglei Gao                                                                                    |      |
|                                                                                                                                                          | 2032 |
| <i>Value Profiles for Encoding Human Variation</i>                                                                                                       |      |
| Taylor Sorensen, Pushkar Mishra, Roma Patel, Michael Henry Tessler, Michiel A. Bakker, Georgina Evans, Iason Gabriel, Noah Goodman and Verena Rieser     |      |
|                                                                                                                                                          | 2048 |
| <i>Language Models as Causal Effect Generators</i>                                                                                                       |      |
| Lucius E.J. Bynum and Kyunghyun Cho                                                                                                                      |      |
|                                                                                                                                                          | 2097 |
| <i>Constructions are Revealed in Word Distributions</i>                                                                                                  |      |
| Joshua Rozner, Leonie Weissweiler, Kyle Mahowald and Cory Shain                                                                                          |      |
|                                                                                                                                                          | 2117 |

|                                                                                                                                                                                          |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>CodeMixBench: Evaluating Code-Mixing Capabilities of LLMs Across 18 Languages</i>                                                                                                     |      |
| Yilun Yang and Yekun Chai .....                                                                                                                                                          | 2140 |
| <i>RBPTool: A Deep Language Model Framework for Multi-Resolution RBP-RNA Binding Prediction and RNA Molecule Design</i>                                                                  |      |
| Jiyue Jiang, Yitao Xu, Zikang Wang, Yihan Ye, Yanruisheng Shao, Yuheng Shan, Jiuming Wang, Xiaodan Fan, Jiao Yuan and Yu Li .....                                                        | 2171 |
| <i>Unveiling Internal Reasoning Modes in LLMs: A Deep Dive into Latent Reasoning vs. Factual Shortcuts with Attribute Rate Ratio</i>                                                     |      |
| Yiran Yang, Haifeng Sun, Jingyu Wang, Qi Qi, Zirui Zhuang, Huazheng Wang, Pengfei Ren, Jing Wang and Jianxin Liao .....                                                                  | 2187 |
| <i>SAE-SSV: Supervised Steering in Sparse Representation Spaces for Reliable Control of Language Models</i>                                                                              |      |
| Zirui He, Mingyu Jin, Bo Shen, Ali Payani, Yongfeng Zhang and Mengnan Du .....                                                                                                           | 2208 |
| <i>BabyLM's First Constructions: Causal interventions provide a signal of learning</i>                                                                                                   |      |
| Joshua Rozner, Leonie Weissweiler and Cory Shain .....                                                                                                                                   | 2238 |
| <i>Effective Red-Teaming of Policy-Adherent Agents</i>                                                                                                                                   |      |
| Itay Nakash, George Kour, Koren Lazar, Matan Vetzler, Guy Uziel and Ateret Anaby Tavor .....                                                                                             | 2251 |
| <i>CondAmbigQA: A Benchmark and Dataset for Conditional Ambiguous Question Answering</i>                                                                                                 |      |
| Zongxi Li, Yang Li, Haoran Xie and S. Joe Qin .....                                                                                                                                      | 2270 |
| <i>SafeScientist: Enhancing AI Scientist Safety for Risk-Aware Scientific Discovery</i>                                                                                                  |      |
| Kunlun Zhu, Jiaxun Zhang, Ziheng Qi, Nuoxing Shang, Zijia Liu, Peixuan Han, Yue Su, Haofei Yu and Jiaxuan You .....                                                                      | 2290 |
| <i>Improving Informally Romanized Language Identification</i>                                                                                                                            |      |
| Adrian Benton, Alexander Gutkin, Christo Kirov and Brian Roark .....                                                                                                                     | 2319 |
| <i>Integral Transformer: Denoising Attention, Not Too Much Not Too Little</i>                                                                                                            |      |
| Ivan Kobyzev, Abbas Ghaddar, Dingtao Hu and Boxing Chen .....                                                                                                                            | 2338 |
| <i>CHENGYU-BENCH: Benchmarking Large Language Models for Chinese Idiom Understanding and Use</i>                                                                                         |      |
| Yicheng Fu, Zhemin Huang, Liuxin Yang, Yumeng Lu and Zhongdongming Dai .....                                                                                                             | 2356 |
| <i>Improving Cross Lingual Transfer by Pretraining with Active Forgetting</i>                                                                                                            |      |
| Divyanshu Aggarwal, Ashutosh Sathe and Sunayana Sitaram .....                                                                                                                            | 2368 |
| <i>Re-Align: Aligning Vision Language Models via Retrieval-Augmented Direct Preference Optimization</i>                                                                                  |      |
| Shuo Xing, Peiran Li, Yuping Wang, Ruizheng Bai, Yueqi Wang, Chan-Wei Hu, Chengxuan Qian, Huaxiu Yao and Zhengzhong Tu .....                                                             | 2380 |
| <i>To Mask or to Mirror: Human-AI Alignment in Collective Reasoning</i>                                                                                                                  |      |
| Crystal Qian, Aaron T Parisi, Cl  mentine Bouleau, Vivian Tsai, Ma  l Lebreton and Lucas Dixon .....                                                                                     | 2399 |
| <i>SWAN: An Efficient and Scalable Approach for Long-Context Language Modeling</i>                                                                                                       |      |
| Krishna C Puvvada, Faisal Ladhak, Santiago Akle Serano, Cheng-Ping Hsieh, Shantanu Acharya, Somshubra Majumdar, Fei Jia, Samuel Krizan, Simeng Sun, Dima Rekesh and Boris Ginsburg ..... | 2425 |
| <i>LLMs Behind the Scenes: Enabling Narrative Scene Illustration</i>                                                                                                                     |      |
| Melissa Roemmele, John Joon Young Chung, Taewook Kim, Yuqian Sun, Alex Calderwood and Max Kreminski .....                                                                                | 2440 |

|                                                                                                                                                                                 |      |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>REARANK: Reasoning Re-ranking Agent via Reinforcement Learning</i>                                                                                                           |      |
| Le Zhang, Bo Wang, Xipeng Qiu, Siva Reddy and Aishwarya Agrawal                                                                                                                 | 2459 |
| <i>Large Language Models Do Multi-Label Classification Differently</i>                                                                                                          |      |
| Marcus Ma, Georgios Chochlakis, Niyantha Maruthu Pandiyan, Jesse Thomason and Shrikanth Narayanan                                                                               | 2473 |
| <i>FilBench: Can LLMs Understand and Generate Filipino?</i>                                                                                                                     |      |
| Lester James Validad Miranda, Elyanah Aco, Conner G. Manuel, Jan Christian Blaise Cruz and Joseph Marvin Imperial                                                               | 2497 |
| <i>M-ABSA: A Multilingual Dataset for Aspect-Based Sentiment Analysis</i>                                                                                                       |      |
| ChengYan Wu, Bolei Ma, Yihong Liu, Zheyu Zhang, Ningyuan Deng, Yanshu Li, Baolan Chen, Yi Zhang, Yun Xue and Barbara Plank                                                      | 2531 |
| <i>RuCCoD: Towards Automated ICD Coding in Russian</i>                                                                                                                          |      |
| Alexandr Nesterov, Andrey Sakhovskiy, Ivan Sviridov, Airat Valiev, Vladimir Makharev, Petr Anokhin, Galina Zubkova and Elena Tutubalina                                         | 2559 |
| <i>Code to Think, Think to Code: A Survey on Code-Enhanced Reasoning and Reasoning-Driven Code Intelligence in LLMs</i>                                                         |      |
| Dayu Yang, Tianyang Liu, Daoan Zhang, Antoine Simoulin, Xiaoyi Liu, Yuwei Cao, Zhaopu Teng, Xin Qian, Grey Yang, Jiebo Luo and Julian McAuley                                   | 2587 |
| <i>Efficient Model Development through Fine-tuning Transfer</i>                                                                                                                 |      |
| Pin-Jie Lin, Rishab Balasubramanian, Fengyuan Liu, Nikhil Kandpal and Tu Vu                                                                                                     | 2618 |
| <i>Language Mixing in Reasoning Language Models: Patterns, Impact, and Internal Causes</i>                                                                                      |      |
| Mingyang Wang, Lukas Lange, Heike Adel, Yunpu Ma, Jannik Strötgen and Hinrich Schuetze                                                                                          | 2638 |
| <i>User Feedback in Human-LLM Dialogues: A Lens to Understand Users But Noisy as a Learning Signal</i>                                                                          |      |
| Yuhan Liu, Michael JQ Zhang and Eunsol Choi                                                                                                                                     | 2667 |
| <i>Read to Hear: A Zero-Shot Pronunciation Assessment Using Textual Descriptions and LLMs</i>                                                                                   |      |
| Yu-Wen Chen, Melody Ma and Julia Hirschberg                                                                                                                                     | 2683 |
| <i>COCO-Tree: Compositional Hierarchical Concept Trees for Enhanced Reasoning in Vision-Language Models</i>                                                                     |      |
| Sanchit Sinha, Guangzhi Xiong and Aidong Zhang                                                                                                                                  | 2696 |
| <i>SurveyGen: Quality-Aware Scientific Survey Generation with Large Language Models</i>                                                                                         |      |
| Tong Bao, Mir Tafseer Nayeem, Davood Rafiei and Chengzhi Zhang                                                                                                                  | 2713 |
| <i>VoiceCraft-X: Unifying Multilingual, Voice-Cloning Speech Synthesis and Speech Editing</i>                                                                                   |      |
| Zhisheng Zheng, Puyuan Peng, Anuj Diwan, Cong Phuoc Huynh, Xiaohang Sun, Zhu Liu, Vimal Bhat and David Harwath                                                                  | 2738 |
| <i>From Generation to Judgment: Opportunities and Challenges of LLM-as-a-judge</i>                                                                                              |      |
| Dawei Li, Bohan Jiang, Liangjie Huang, Alimohammad Beigi, Chengshuai Zhao, Zhen Tan, Amrita Bhattacharjee, Yuxuan Jiang, Canyu Chen, Tianhao Wu, Kai Shu, Lu Cheng and Huan Liu | 2758 |
| <i>MultiMatch: Multihead Consistency Regularization Matching for Semi-Supervised Text Classification</i>                                                                        |      |
| Iustin Sirbu, Robert-Adrian Popovici, Cornelia Caragea, Stefan Trausan-Matu and Traian Rebedea                                                                                  | 2793 |

|                                                                                                                                                                                                                                            |      |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>TTT-Bench: A Benchmark for Evaluating Reasoning Ability with Simple and Novel Tic-Tac-Toe-style Games</i>                                                                                                                               |      |
| Prakamya Mishra, Jiang Liu, Jialian Wu, Xiaodong Yu, Zicheng Liu and Emad Barsoum . . .                                                                                                                                                    | 2810 |
| <i>Learning from Diverse Reasoning Paths with Routing and Collaboration</i>                                                                                                                                                                |      |
| Zhenyu Lei, Zhen Tan, Song Wang, Yaochen Zhu, Zihan Chen, Yushun Dong and Jundong Li                                                                                                                                                       | 2833 |
| <i>Ask Patients with Patience: Enabling LLMs for Human-Centric Medical Dialogue with Grounded Reasoning</i>                                                                                                                                |      |
| Jiayuan Zhu, Jiazhen Pan, Yuyuan Liu, Fenglin Liu and Junde Wu . . . . .                                                                                                                                                                   | 2847 |
| <i>MedHallu: A Comprehensive Benchmark for Detecting Medical Hallucinations in Large Language Models</i>                                                                                                                                   |      |
| Shrey Pandit, Jiawei Xu, Junyuan Hong, Zhangyang Wang, Tianlong Chen, Kaidi Xu and Ying Ding . . . . .                                                                                                                                     | 2859 |
| <i>NUTMEG: Separating Signal From Noise in Annotator Disagreement</i>                                                                                                                                                                      |      |
| Jonathan Ivey, Susan Gauch and David Jurgens . . . . .                                                                                                                                                                                     | 2875 |
| <i>Alignment Quality Index (AQI) : Beyond Refusals: AQI as an Intrinsic Alignment Diagnostic via Latent Geometry, Cluster Divergence, and Layer wise Pooled Representations</i>                                                            |      |
| Abhilekh Borah, Chhavi Sharma, Danush Khanna, Utkarsh Bhatt, Gurpreet Singh, Hasnat Md Abdullah, Raghav Kaushik Ravi, Vinija Jain, Jyoti Patel, Shubham Singh, Vasu Sharma, Arpita Vats, Rahul Raja, Aman Chadha and Amitava Das . . . . . | 2889 |
| <i>MythTriage: Scalable Detection of Opioid Use Disorder Myths on a Video-Sharing Platform</i>                                                                                                                                             |      |
| Hayoung Jung, Shravika Mittal, Ananya Aatreya, Navreet Kaur, Munmun De Choudhury and Tanu Mitra . . . . .                                                                                                                                  | 2949 |
| <i>Demystifying optimized prompts in language models</i>                                                                                                                                                                                   |      |
| Rimon Melamed, Lucas Hurley McCabe and H Howie Huang . . . . .                                                                                                                                                                             | 2984 |
| <i>Whisper-UT: A Unified Translation Framework for Speech and Text</i>                                                                                                                                                                     |      |
| Cihan Xiao, Matthew Wiesner, Debashish Chakraborty, Reno Kriz, Keith Cunningham, Kenton Murray, Kevin Duh, Luis Tavaréz-Arce, Paul McNamee and Sanjeev Khudanpur . . . . .                                                                 | 3001 |
| <i>Unleashing the Reasoning Potential of LLMs by Critique Fine-Tuning on One Problem</i>                                                                                                                                                   |      |
| Yubo Wang, Ping Nie, Kai Zou, Lijun Wu and Wenhui Chen . . . . .                                                                                                                                                                           | 3018 |
| <i>Active Layer-Contrastive Decoding Reduces Hallucination in Large Language Model Generation</i>                                                                                                                                          |      |
| Hongxiang Zhang, Hao Chen, Muhao Chen and Tianyi Zhang . . . . .                                                                                                                                                                           | 3029 |
| <i>BBScoreV2: Learning Time-Evolution and Latent Alignment from Stochastic Representation</i>                                                                                                                                              |      |
| Tianhao Zhang, Zhecheng Sheng, Zhexiao Lin, Chen Jiang and Dongyeop Kang . . . . .                                                                                                                                                         | 3048 |
| <i>SAND: Boosting LLM Agents with Self-Taught Action Deliberation</i>                                                                                                                                                                      |      |
| Yu Xia, Yiran Jenny Shen, Junda Wu, Tong Yu, Sungchul Kim, Ryan A. Rossi, Lina Yao and Julian McAuley . . . . .                                                                                                                            | 3063 |
| <i>LLMs as World Models: Data-Driven and Human-Centered Pre-Event Simulation for Disaster Impact Assessment</i>                                                                                                                            |      |
| Lingyao Li, Dawei Li, Zhenhui Ou, Xiaoran Xu, Jingxiao Liu, Zihui Ma, Runlong Yu and Min Deng . . . . .                                                                                                                                    | 3079 |
| <i>Mind the Value-Action Gap: Do LLMs Act in Alignment with Their Values?</i>                                                                                                                                                              |      |
| Hua Shen, Nicholas Clark and Tanu Mitra . . . . .                                                                                                                                                                                          | 3098 |

|                                                                                                                                            |      |
|--------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Two Heads Are Better Than One: Dual-Model Verbal Reflection at Inference-Time</i>                                                       |      |
| Jiazheng Li, Yuxiang Zhou, Junru Lu, Gladys Tyen, Lin Gui, Cesare Aloisi and Yulan He . .                                                  | 3120 |
| <i>Image Embedding Sampling Method for Diverse Captioning</i>                                                                              |      |
| Sania Waheed and Na Min An . . . . .                                                                                                       | 3142 |
| <i>Diagnosing Memorization in Chain-of-Thought Reasoning, One Token at a Time</i>                                                          |      |
| Huihan Li, You Chen, Siyuan Wang, Yixin He, Ninareh Mehrabi, Rahul Gupta and Xiang Ren                                                     | 3159 |
| <i>FANS: Formal Answer Selection for LLM Natural Language Math Reasoning Using Lean4</i>                                                   |      |
| Jiarui Yao, Ruida Wang and Tong Zhang . . . . .                                                                                            | 3182 |
| <i>Date Fragments: A Hidden Bottleneck of Tokenization for Temporal Reasoning</i>                                                          |      |
| Gagan Bhatia, Maxime Peyrard and Wei Zhao . . . . .                                                                                        | 3202 |
| <i>Measuring Risk of Bias in Biomedical Reports: The RoBBR Benchmark</i>                                                                   |      |
| Jianyou Wang, Weili Cao, Longtian Bao, Youze Zheng, Gil Pasternak, Kaicheng Wang, Xiaoyue Wang, Ramamohan Paturi and Leon Bergen . . . . . | 3221 |
| <i>SHIFT: Selected Helpful Informative Frame for Video-guided Machine Translation</i>                                                      |      |
| Boyu Guan, Chuang Han, Yining Zhang, Yupu Liang, Zhiyang Zhang, Yang Zhao and Chengqing Zong . . . . .                                     | 3250 |
| <i>Surge: On the Potential of Large Language Models as General-Purpose Surrogate Code Executors</i>                                        |      |
| Bohan Lyu, Siqiao Huang, Zichen Liang, Qian Sun and Jiaming Zhang . . . . .                                                                | 3269 |
| <i>Few-Shot Learning Translation from New Languages</i>                                                                                    |      |
| Carlos Mullov and Alexander Waibel . . . . .                                                                                               | 3310 |
| <i>Humanizing Machines: Rethinking LLM Anthropomorphism Through a Multi-Level Framework of Design</i>                                      |      |
| Yunze Xiao, Lynnette Hui Xian Ng, Jiarui Liu and Mona T. Diab . . . . .                                                                    | 3332 |
| <i>TokenSkip: Controllable Chain-of-Thought Compression in LLMs</i>                                                                        |      |
| Heming Xia, Chak Tou Leong, Wenjie Wang, Yongqi Li and Wenjie Li . . . . .                                                                 | 3352 |
| <i>Are Generative Models Underconfident? Better Quality Estimation with Boosted Model Probability</i>                                      |      |
| Tu Anh Dinh and Jan Niehues . . . . .                                                                                                      | 3365 |
| <i>reWordBench: Benchmarking and Improving the Robustness of Reward Models with Transformed Inputs</i>                                     |      |
| Zhaofeng Wu, Michihiro Yasunaga, Andrew Cohen, Yoon Kim, Asli Celikyilmaz and Marjan Ghazvininejad . . . . .                               | 3384 |
| <i>Why Do Some Inputs Break Low-Bit LLM Quantization?</i>                                                                                  |      |
| Ting-Yun Chang, Muru Zhang, Jesse Thomason and Robin Jia . . . . .                                                                         | 3411 |
| <i>LiteASR: Efficient Automatic Speech Recognition with Low-Rank Approximation</i>                                                         |      |
| Keisuke Kamahori, Jungo Kasai, Noriyuki Kojima and Baris Kasikci . . . . .                                                                 | 3431 |
| <i>AROMA: Autonomous Rank-one Matrix Adaptation</i>                                                                                        |      |
| Hao Nan Sheng, Zhi-Yong Wang, Hing Cheung So and Mingrui Yang . . . . .                                                                    | 3444 |
| <i>Large Language Models Have Intrinsic Meta-Cognition, but Need a Good Lens</i>                                                           |      |
| Ziyang Ma, Qingyue Yuan, Zhenglin Wang and Deyu Zhou . . . . .                                                                             | 3461 |

|                                                                                                                                                                             |      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Anchoring-Guidance Fine-Tuning (AnGFT): Elevating Professional Response Quality in Role-Playing Conversational Agents</i>                                                |      |
| Qibin Li, Zhen Xu, Shengyuan Bai, Nianmin Yao, Kaili Sun, Bowen Wu, Ying Li and Baoxun Wang .....                                                                           | 3479 |
| <i>RiTTA: Modeling Event Relations in Text-to-Audio Generation</i>                                                                                                          |      |
| Yuhang He, Yash Jain, Xubo Liu, Andrew Markham and Vibhav Vineet .....                                                                                                      | 3498 |
| <i>Shallow Focus, Deep Fixes: Enhancing Shallow Layers Vision Attention Sinks to Alleviate Hallucination in LVLMS</i>                                                       |      |
| Xiaofeng Zhang, Yihao Quan, Chen Shen, Chaochen Gu, Xiaosong Yuan, Shaotian Yan, Jiawei Cao, Hao Cheng, Kaijie Wu and Jieping Ye .....                                      | 3513 |
| <i>WangchanThaiInstruct: An instruction-following Dataset for Culture-Aware, Multitask, and Multi-domain Evaluation in Thai</i>                                             |      |
| Peerat Limkonchotiwat, Pume Tuchinda, Lalita Lowphansirikul, Surapon Nonesung, Panuthep Tasawong, Alham Fikri Aji, Can Udomcharoenchaikit and Sarana Nutanong .....         | 3536 |
| <i>MemeReaCon: Probing Contextual Meme Understanding in Large Vision-Language Models</i>                                                                                    |      |
| Zhengyi Zhao, Shubo Zhang, Yuxi Zhang, Yanxi Zhao, Yifan Zhang, Zezhong Wang, Huimin Wang, Yutian Zhao, Bin Liang, Yefeng Zheng, Binyang Li, Kam-Fai Wong and Xian Wu ..... | 3560 |
| <i>A Comprehensive Literary Chinese Reading Comprehension Dataset with an Evidence Curation Based Solution</i>                                                              |      |
| Dongning Rao, Rongchu Zhou, Peng Chen and Zhihua Jiang .....                                                                                                                | 3584 |
| <i>Dialect-SQL: An Adaptive Framework for Bridging the Dialect Gap in Text-to-SQL</i>                                                                                       |      |
| Jie Shi, Xi Cao, Bo Xu, Jiaqing Liang, Yanghua Xiao, Jia Chen, Peng Wang and Wei Wang .....                                                                                 | 3605 |
| <i>FinMTEB: Finance Massive Text Embedding Benchmark</i>                                                                                                                    |      |
| Yixuan Tang and Yi Yang .....                                                                                                                                               | 3621 |
| <i>Scaling Rich Style-Prompted Text-to-Speech Datasets</i>                                                                                                                  |      |
| Anuj Diwan, Zhisheng Zheng, David Harwath and Eunsol Choi .....                                                                                                             | 3640 |
| <i>Exploring Changes in Nation Perception with Nationality-Assigned Personas in LLMs</i>                                                                                    |      |
| Mahammed Kamruzzaman and Gene Louis Kim .....                                                                                                                               | 3661 |
| <i>Eliciting Implicit Acoustic Styles from Open-domain Instructions to Facilitate Fine-grained Controllable Generation of Speech</i>                                        |      |
| Jianxing Yu, Gou Zihao, Chen Li, Zhisheng Wang, Peiji Yang, Wenqing Chen and Jian Yin .....                                                                                 | 3680 |
| <i>OBLIVATE: Robust and Practical Machine Unlearning for Large Language Models</i>                                                                                          |      |
| Xiaoyu Xu, Minxin Du, Qingqing Ye and Haibo Hu .....                                                                                                                        | 3697 |
| <i>AdaptThink: Reasoning Models Can Learn When to Think</i>                                                                                                                 |      |
| Jiajie Zhang, Nianyi Lin, Lei Hou, Ling Feng and Juanzi Li .....                                                                                                            | 3717 |
| <i>T<sup>2</sup>: An Adaptive Test-Time Scaling Strategy for Contextual Question Answering</i>                                                                              |      |
| Zhengyi Zhao, Shubo Zhang, Zezhong Wang, Huimin Wang, Yutian Zhao, Bin Liang, Yefeng Zheng, Binyang Li, Kam-Fai Wong and Xian Wu .....                                      | 3732 |
| <i>Non-Existent Relationship: Fact-Aware Multi-Level Machine-Generated Text Detection</i>                                                                                   |      |
| Yang Wu, Ruijia Wang and Jie Wu .....                                                                                                                                       | 3758 |
| <i>Calibrating Verbal Uncertainty as a Linear Feature to Reduce Hallucinations</i>                                                                                          |      |
| Ziwei Ji, Lei Yu, Yeskendir Koishchenov, Yejin Bang, Anthony Hartshorn, Alan Schelten, Cheng Zhang, Pascale Fung and Nicola Cancedda .....                                  | 3770 |



|                                                                                                                                                    |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>JUREX-4E: Juridical Expert-Annotated Four-Element Knowledge Base for Legal Reasoning</i>                                                        |      |
| Huanghai Liu, Quzhe Huang, Qingjing Chen, Yiran HU, Jiayu Ma, Yun Liu, Weixing Shen and Yansong Feng .....                                         | 3795 |
| <i>CIE: Controlling Language Model Text Generations Using Continuous Signals</i>                                                                   |      |
| Vinay Samuel, Harshita Diddee, Yiming Zhang and Daphne Ippolito .....                                                                              | 3816 |
| <i>Stand on The Shoulders of Giants: Building JailExpert from Previous Attack Experience</i>                                                       |      |
| Xi Wang, Songlei Jian, Shasha Li, Xiaopeng Li, Bin Ji, Ma Jun, Xiaodong Liu, Jing Wang, Jianfeng Zhang, Jie Yu, Feilong Bao and Wangbaosheng ..... | 3827 |
| <i>Language-to-Space Programming for Training-Free 3D Visual Grounding</i>                                                                         |      |
| Boyu Mi, Hanqing Wang, Tai Wang, Yilun Chen and Jiangmiao Pang .....                                                                               | 3845 |
| <i>RAG-Instruct: Boosting LLMs with Diverse Retrieval-Augmented Instructions</i>                                                                   |      |
| Wanlong Liu, Junying Chen, Ke Ji, Li Zhou, Wenyu Chen and Benyou Wang .....                                                                        | 3866 |
| <i>AdaRewriter: Unleashing the Power of Prompting-based Conversational Query Reformulation via Test-Time Adaptation</i>                            |      |
| Yilong Lai, Jialong Wu, Zhenglin Wang and Deyu Zhou .....                                                                                          | 3890 |
| <i>SmartBench: Is Your LLM Truly a Good Chinese Smartphone Assistant?</i>                                                                          |      |
| Xudong Lu, Haohao Gao, Renshou Wu, Shuai Ren, Xiaoxin Chen, Hongsheng Li and Fangyuan Li .....                                                     | 3907 |
| <i>F2TEval: Human-Aligned Multi-Dimensional Evaluation for Figure-to-Text Task</i>                                                                 |      |
| Tan Yue, Rui Mao, Zilong Song, Zonghai Hu and Dongyan Zhao .....                                                                                   | 3933 |
| <i>Icon<sup>2</sup>: Aligning Large Language Models Using Self-Synthetic Preference Data via Inherent Regulation</i>                               |      |
| Qiyuan Chen, Hongsen Huang, Qian Shao, Jiahe Chen, Jintai Chen, Hongxia Xu, Renjie Hua, Ren Chuan and Jian Wu .....                                | 3950 |
| <i>DSCD: Large Language Model Detoxification with Self-Constrained Decoding</i>                                                                    |      |
| Ming Dong, Jinkui Zhang, Bolong Zheng, Xinhui Tu, Po Hu and Tingting He .....                                                                      | 3970 |
| <i>From Reasoning to Answer: Empirical, Attention-Based and Mechanistic Insights into Distilled Deep-Seek R1 Models</i>                            |      |
| Jue Zhang, Qingwei Lin, Saravan Rajmohan and Dongmei Zhang .....                                                                                   | 3986 |
| <i>Quantifying Language Disparities in Multilingual Large Language Models</i>                                                                      |      |
| Songbo Hu, Ivan Vulić and Anna Korhonen .....                                                                                                      | 4004 |
| <i>KoBLEX: Open Legal Question Answering with Multi-hop Reasoning</i>                                                                              |      |
| Jihyung Lee, Daehui Kim, Seonjeong Hwang, Hyoungun Kim and Gary Lee .....                                                                          | 4020 |
| <i>End-to-End Learnable Psychiatric Scale Guided Risky Post Screening for Depression Detection on Social Media</i>                                 |      |
| Bichen Wang, Yuzhe Zi, Yixin Sun, Hao Yang, Yanyan Zhao and Bing Qin .....                                                                         | 4055 |
| <i>ReAgent: Reversible Multi-Agent Reasoning for Knowledge-Enhanced Multi-Hop QA</i>                                                               |      |
| Zhao Xinjie, Fan Gao, Xingyu Song, Yingjian Chen, Rui Yang, Yanran Fu, Yuyang Wang, Yusuke Iwasawa, Yutaka Matsuo and Irene Li .....               | 4068 |
| <i>Matter-of-Fact: A Benchmark for Verifying the Feasibility of Literature-Supported Claims in Materials Science</i>                               |      |
| Peter Jansen, Samiah Hassan and Ruoyao Wang .....                                                                                                  | 4091 |

|                                                                                                                                                              |      |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>ModRWKV: Transformer Multimodality in Linear Time</i>                                                                                                     |      |
| Jiale Kang, Ziyin Yue, Qingyu Yin, Rui Jiang, Weile Li, Zening Lu and Zhouan Ji                                                                              | 4104 |
| <i>Multimedia Event Extraction with LLM Knowledge Editing</i>                                                                                                |      |
| Jiaao Yu, Yijing Lin, Zhipeng Gao, Xuesong Qiu and Lanlan Rui                                                                                                | 4117 |
| <i>Exploring the Impact of Personality Traits on LLM Toxicity and Bias</i>                                                                                   |      |
| Shuo Wang, Renhao Li, Xi Chen, Yulin Yuan, Min Yang and Derek F. Wong                                                                                        | 4126 |
| <i>Task-aware Contrastive Mixture of Experts for Quadruple Extraction in Conversations with Code-like Replies and Non-opinion Detection</i>                  |      |
| Chenyuan He, Yuxiang Jia, Fei Gao, Senbin Zhu, Hongde Liu, Hongying Zan and Min Peng                                                                         | 4145 |
| <i>Mitigating Biases in Language Models via Bias Unlearning</i>                                                                                              |      |
| Dianqing Liu, Yi Liu, Guoqing Jin and Zhendong Mao                                                                                                           | 4161 |
| <i>UNComp: Can Matrix Entropy Uncover Sparsity? — A Compressor Design from an Uncertainty-Aware Perspective</i>                                              |      |
| Jing Xiong, Jianghan Shen, Fanghua Ye, Chaofan Tao, Zhongwei Wan, Jianqiao Lu, Xun Wu, Chuanyang Zheng, Zhijiang Guo, Min Yang, Lingpeng Kong and Ngai Wong  | 4180 |
| <i>Superpose Task-specific Features for Model Merging</i>                                                                                                    |      |
| Haiquan Qiu, You Wu, Dong Li, Jianmin Guo and Quanming Yao                                                                                                   | 4201 |
| <i>FinRAGBench-V: A Benchmark for Multimodal RAG with Visual Citation in the Financial Domain</i>                                                            |      |
| Suifeng Zhao, Zhuoran Jin, Sujian Li and Jun Gao                                                                                                             | 4216 |
| <i>BacktrackAgent: Enhancing GUI Agent with Error Detection and Backtracking Mechanism</i>                                                                   |      |
| Qinzhao Wu, Pengzhi Gao, Wei Liu and Jian Luan                                                                                                               | 4251 |
| <i>Diffusion vs. Autoregressive Language Models: A Text Embedding Perspective</i>                                                                            |      |
| Siyue Zhang, Yilun Zhao, Liyuan Geng, Arman Cohan, Anh Tuan Luu and Chen Zhao                                                                                | 4274 |
| <i>BannerAgency: Advertising Banner Design with Multimodal LLM Agents</i>                                                                                    |      |
| Heng Wang, Yotaro Shimose and Shingo Takamatsu                                                                                                               | 4305 |
| <i>DIDS: Domain Impact-aware Data Sampling for Large Language Model Training</i>                                                                             |      |
| Weijie Shi, Jipeng Zhang, Yaguang Wu, Jingzhi Fang, Shibo Zhang, Yao Zhao, Hao Chen, Ruiyuan Zhang, Yue Cui, Jia Zhu, Sirui Han, Jiajie Xu and Xiaofang Zhou | 4331 |
| <i>Training LLMs to be Better Text Embedders through Bidirectional Reconstruction</i>                                                                        |      |
| Chang Su, Dengliang Shi, Siyuan Huang, Jintao Du, Changhua Meng, Yu Cheng, Weiqiang Wang and Zhouhan Lin                                                     | 4352 |
| <i>ReMedy: Learning Machine Translation Evaluation from Human Preferences with Reward Modeling</i>                                                           |      |
| Shaomu Tan and Christof Monz                                                                                                                                 | 4371 |
| <i>SolEval: Benchmarking Large Language Models for Repository-level Solidity Smart Contract Generation</i>                                                   |      |
| Zhiyuan Peng, Xin Yin, Rui Qian, Peiqin Lin, YongKang Liu, Hao Zhang, Chenhao Ying and Yuan Luo                                                              | 4389 |
| <i>In-Context Learning Boosts Speech Recognition via Human-like Adaptation to Speakers and Language Varieties</i>                                            |      |
| Nathan Roll, Calbert Graham, Yuka Tatsumi, Kim Tien Nguyen, Meghan Sumner and Dan Jurafsky                                                                   | 4413 |



|                                                                                                                                                                                                                                        |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Reasoning Model Unlearning: Forgetting Traces, Not Just Answers, While Preserving Reasoning Skills</i><br>Changsheng Wang, Chongyu Fan, Yihua Zhang, Jinghan Jia, Dennis Wei, Parikshit Ram, Nathalie Baracaldo and Sijia Liu ..... | 4428 |
| <i>Chain-of-Talkers (CoTalk): Fast Human Annotation of Dense Image Captions</i><br>Yijun Shen, Delong Chen, Fan Liu, Xingyu Wang, Chuanyi Zhang, Liang Yao and Yuhui Zheng                                                             | 4445 |
| <i>DecoupleSearch: Decouple Planning and Search via Hierarchical Reward Modeling</i><br>Hao Sun, Zile Qiao, Bo Wang, Guoxin Chen, Yingyan Hou, Yong Jiang, Pengjun Xie, Fei Huang and Yan Zhang .....                                  | 4466 |
| <i>RewardDS: Privacy-Preserving Fine-Tuning for Large Language Models via Reward Driven Data Synthesis</i><br>Jianwei Wang, Chengming Shi, Junyao Yang, Haoran Li, Qianli Ma, Huiping Zhuang, Cen Chen and Ziqian Zeng .....           | 4480 |
| <i>Synergizing Multimodal Temporal Knowledge Graphs and Large Language Models for Social Relation Recognition</i><br>Haorui Wang, Zheng Wang, Yuxuan Zhang, Bo Wang and Bin Wu .....                                                   | 4502 |
| <i>LegalSearchLM: Rethinking Legal Case Retrieval as Legal Elements Generation</i><br>Chaeun Kim, Jinu Lee and Wonseok Hwang .....                                                                                                     | 4522 |
| <i>ChartMind: A Comprehensive Benchmark for Complex Real-world Multimodal Chart Question Answering</i><br>Jingxuan Wei, Nan Xu, Junnan Zhu, Haoyanni, Gaowei Wu, Qi Chen, Bihui Yu and Lei Wang                                        | 4556 |
| <i>COLA: Collaborative Multi-Agent Framework with Dynamic Task Scheduling for GUI Automation</i><br>Di Zhao, Longhui Ma, Siwei Wang, Miao Wang and Zhao Lv .....                                                                       | 4571 |
| <i>DASA-Trans-STM: Adaptive Efficient Transformer for Short Text Matching using Data Augmentation and Semantic Awareness</i><br>Jiguo Liu, Chao Liu, Meimei Li, Nan Li, Shihao Gao and Dali Zhu .....                                  | 4595 |
| <i>Pruning the Paradox: How CLIP's Most Informative Heads Enhance Performance While Amplifying Bias</i><br>Avinash Madasu, Vasudev Lal and Phillip Howard .....                                                                        | 4612 |
| <i>CoLA: Compute-Efficient Pre-Training of LLMs via Low-Rank Activation</i><br>Ziyue Liu, Ruijie Zhang, Zhengyang Wang, Mingsong Yan, Zi Yang, Paul D. Hovland, Bogdan Nicolae, Franck Cappello, Sui Tang and Zheng Zhang .....        | 4628 |
| <i>TS-CLIP: Time Series Understanding by CLIP</i><br>Ziwen Chen, Xiaoyuan Zhang and Ming Zhu .....                                                                                                                                     | 4647 |
| <i>MultiAgentESC: A LLM-based Multi-Agent Collaboration Framework for Emotional Support Conversation</i><br>Yangyang Xu, Jinpeng Hu, Zhuoer Zhao, Zhangling Duan, Xiao Sun and Xun Yang .....                                          | 4666 |
| <i>Continuously Steering LLMs Sensitivity to Contextual Knowledge with Proxy Models</i><br>Yilin Wang, Heng Wang, Yuyang Bai and Minnan Luo .....                                                                                      | 4683 |
| <i>Probing LLM World Models: Enhancing Guesstimation with Wisdom of Crowds Decoding</i><br>Yun-Shiuan Chuang, Sameer Narendran, Nikunj Harlalka, Alexander Cheung, Sizhe Gao, Siddharth Suresh, Junjie Hu and Timothy T. Rogers .....  | 4700 |

|                                                                                                                                                  |      |
|--------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Recall with Reasoning: Chain-of-Thought Distillation for Mamba’s Long-Context Memory and Extrapolation</i>                                    |      |
| Jun-Yu Ma, Tianqing Fang, Zhisong Zhang, Hongming Zhang, Haitao Mi and Dong Yu . . .                                                             | 4715 |
| <i>Scalable Data Synthesis through Human-like Cognitive Imitation and Data Recombination</i>                                                     |      |
| Zhongyi Ye, Weitai Zhang, Xinyuan Zhou, Yongxin Zhu, Ninghui Rao and Enhong Chen . .                                                             | 4722 |
| <i>BeSimulator: A Large Language Model Powered Text-based Behavior Simulator</i>                                                                 |      |
| Jianan Wang, Bin Li, Jingtao Qi, Xueying Wang, Fu Li and Lihanxun . . . . .                                                                      | 4737 |
| <i>Too Consistent to Detect: A Study of Self-Consistent Errors in LLMs</i>                                                                       |      |
| Hexiang Tan, Fei Sun, Sha Liu, Du Su, Qi Cao, Xin Chen, Jingang Wang, Xunliang Cai, Yuanzhuo Wang, Huawei Shen and Xueqi Cheng . . . . .         | 4756 |
| <i>pFedGPT: Hierarchically Optimizing LoRA Aggregation Weights for Personalized Federated GPT Models</i>                                         |      |
| Zhanming Shen, Tianqi Xu, Hao Wang, Jian Li and Miao Pan . . . . .                                                                               | 4767 |
| <i>QSpec: Speculative Decoding with Complementary Quantization Schemes</i>                                                                       |      |
| Juntao Zhao, Wenhao Lu, Sheng Wang, Lingpeng Kong and Chuan Wu . . . . .                                                                         | 4780 |
| <i>Co-Evolving LLMs and Embedding Models via Density-Guided Preference Optimization for Text Clustering</i>                                      |      |
| Zetong Li, Qinliang Su, Minhua Huang and Yin Yang . . . . .                                                                                      | 4797 |
| <i>P-MMEval: A Parallel Multilingual Multitask Benchmark for Consistent Evaluation of LLMs</i>                                                   |      |
| Yidan Zhang, Yu Wan, Boyi Deng, Baosong Yang, Hao-Ran Wei, Fei Huang, Bowen Yu, Dayi-heng Liu, Junyang Lin, Fei Huang and Jingren Zhou . . . . . | 4810 |
| <i>Single LLM, Multiple Roles: A Unified Retrieval-Augmented Generation Framework Using Role-Specific Token Optimization</i>                     |      |
| Yutao Zhu, Jiajie Jin, Hongjin Qian, Zheng Liu, Zhicheng Dou and Ji-Rong Wen . . . . .                                                           | 4838 |
| <i>TrInk: Ink Generation with Transformer Network</i>                                                                                            |      |
| Zezhong Jin, Shubhang Desai, Xu Chen, Biyi Fang, Zhuoyi Huang, Zhe LI, Chong-Xin Gan, Xiao Tu, Man-Wai Mak, Yan Lu and Shujie Liu . . . . .      | 4858 |
| <i>CalligraphicOCR for Chinese Calligraphy Recognition</i>                                                                                       |      |
| Xiaoyi Bao, Zhongqing Wang, Jinghang Gu and Chu-Ren Huang . . . . .                                                                              | 4866 |
| <i>When Audio and Text Disagree: Revealing Text Bias in Large Audio-Language Models</i>                                                          |      |
| Cheng Wang, Gelei Deng, Xianglin Yang, Han Qiu and Tianwei Zhang . . . . .                                                                       | 4879 |
| <i>RESF: Regularized-Entropy-Sensitive Fingerprinting for Black-Box Tamper Detection of Large Language Models</i>                                |      |
| Pingyi Hu, Xiaofan Bai, Xiaojing Ma, Chaoxiang He, Dongmei Zhang and Bin Benjamin Zhu                                                            | 4890 |
| <i>Model-based Large Language Model Customization as Service</i>                                                                                 |      |
| Zhaomin Wu, Jizhou Guo, Junyi Hou, Bingsheng He, Lixin Fan and Qiang Yang . . . . .                                                              | 4905 |
| <i>Collab-Overcooked: Benchmarking and Evaluating Large Language Models as Collaborative Agents</i>                                              |      |
| Haochen Sun, Shuwen Zhang, Lujie Niu, Lei Ren, Hao Xu, Hao Fu, Fangkun Zhao, Caixia Yuan and Xiaojie Wang . . . . .                              | 4923 |
| <i>Improving Reasoning Capabilities in Small Models through Mixture-of-layers Distillation with Stepwise Attention on Key Information</i>        |      |
| Yao Chen, Jiawei Sheng, Wenyuan Zhang and Tingwen Liu . . . . .                                                                                  | 4953 |

|                                                                                                                                                                                                                         |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Through the Valley: Path to Effective Long CoT Training for Small Language Models</i>                                                                                                                                |      |
| Renjie Luo, Jiaxi Li, Chen Huang and Wei Lu .....                                                                                                                                                                       | 4973 |
| <i>RED: Unleashing Token-Level Rewards from Holistic Feedback via Reward Redistribution</i>                                                                                                                             |      |
| Jiahui Li, Lin Li, Tai-Wei Chang, Kun Kuang, Long Chen, Jun Zhou and Cheng Yang .....                                                                                                                                   | 4994 |
| <i>SDGO: Self-Discrimination-Guided Optimization for Consistent Safety in Large Language Models</i>                                                                                                                     |      |
| Peng Ding, Wen Sun, Dailin Li, Wei Zou, Jiaming Wang, Jiajun Chen and Shujian Huang ..                                                                                                                                  | 5024 |
| <i>InMind: Evaluating LLMs in Capturing and Applying Individual Human Reasoning Styles</i>                                                                                                                              |      |
| Zizhen Li, Chuanhao Li, Yibin Wang, Qi Chen, Diping Song, Yukang Feng, Jianwen Sun, Jiaxin Ai, Fanrui Zhang, Mingzhu Sun and Kaipeng Zhang .....                                                                        | 5039 |
| <i>MIO: A Foundation Model on Multimodal Tokens</i>                                                                                                                                                                     |      |
| Zekun Moore Wang, King Zhu, Chunpu Xu, Wangchunshu Zhou, Jiaheng Liu, Yibo Zhang, Jessie Wang, Ning Shi, Siyu Li, Yizhi Li, Haoran Que, Zhaoxiang Zhang, Yuanxing Zhang, Ge Zhang, Ke Xu, Jie Fu and Wenhao Huang ..... | 5078 |
| <i>DART: Distilling Autoregressive Reasoning to Silent Thought</i>                                                                                                                                                      |      |
| Nan Jiang, Ziming Wu, De-Chuan Zhan, Fuming Lai and Shaobing Lian .....                                                                                                                                                 | 5101 |
| <i>LeTS: Learning to Think-and-Search via Process-and-Outcome Reward Hybridization</i>                                                                                                                                  |      |
| Qi Zhang, Shouqing Yang, Lirong Gao, Hao Chen, Xiaomeng Hu, Jinglei Chen, Jiexiang Wang, Sheng Guo, Bo Zheng, Haobo Wang and Junbo Zhao .....                                                                           | 5110 |
| <i>CYCLE-INSTRUCT: Fully Seed-Free Instruction Tuning via Dual Self-Training and Cycle Consistency</i>                                                                                                                  |      |
| Zhanming Shen, Hao Chen, Yulei Tang, Shaolin Zhu, Wentao Ye, Xiaomeng Hu, Haobo Wang, Gang Chen and Junbo Zhao .....                                                                                                    | 5124 |
| <i>Good Intentions Beyond ACL: Who Does NLP for Social Good, and Where?</i>                                                                                                                                             |      |
| Grace LeFevre, Qingcheng Zeng, Adam Leif, Jason Jewell, Denis Peskoff and Rob Voigt...                                                                                                                                  | 5139 |
| <i>From General Reward to Targeted Reward: Improving Open-ended Long-context Generation Models</i>                                                                                                                      |      |
| Zhihan Guo, Jiele Wu, Wenqian Cui, Yifei Zhang, Minda Hu, Yufei Wang and Irwin King .                                                                                                                                   | 5152 |
| <i>Think in Safety: Unveiling and Mitigating Safety Alignment Collapse in Multimodal Large Reasoning Model</i>                                                                                                          |      |
| Xinyue Lou, You Li, Jinan Xu, Xiangyu Shi, Chi Chen and Kaiyu Huang .....                                                                                                                                               | 5168 |
| <i>Understanding the Modality Gap: An Empirical Study on the Speech-Text Alignment Mechanism of Large Speech Language Models</i>                                                                                        |      |
| Bajian Xiang, Shuaijiang Zhao, Tingwei Guo and Wei Zou .....                                                                                                                                                            | 5188 |
| <i>AssoCiAm: A Benchmark for Evaluating Association Thinking while Circumventing Ambiguity</i>                                                                                                                          |      |
| Yifan Liu, Wenkuan Zhao, Shanshan Zhong, Jinghui Qin, Mingfu Liang, Zhongzhan Huang and Wushao Wen .....                                                                                                                | 5204 |
| <i>M-BRE: Discovering Training Samples for Relation Extraction from Unlabeled Texts with Large Language Models</i>                                                                                                      |      |
| Zexuan Li, Hongliang Dai and Piji Li .....                                                                                                                                                                              | 5221 |
| <i>R-TOFU: Unlearning in Large Reasoning Models</i>                                                                                                                                                                     |      |
| Sangyeon Yoon, Wonje Jeung and Albert No .....                                                                                                                                                                          | 5240 |
| <i>Chat-Driven Text Generation and Interaction for Person Retrieval</i>                                                                                                                                                 |      |
| Zequan Xie, Chuxin Wang, Yejiang Wang, Sihang Cai, Shulei Wang and Tao Jin .....                                                                                                                                        | 5260 |

|                                                                                                                                                                                                            |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Spontaneous Giving and Calculated Greed in Language Models</i>                                                                                                                                          |      |
| Yuxuan Li and Hirokazu Shirado .....                                                                                                                                                                       | 5272 |
| <i>SenDetEX: Sentence-Level AI-Generated Text Detection for Human-AI Hybrid Content via Style and Context Fusion</i>                                                                                       |      |
| Lei Jiang, Desheng Wu and Xiaolong Zheng .....                                                                                                                                                             | 5288 |
| <i>Judge and Improve: Towards a Better Reasoning of Knowledge Graphs with Large Language Models</i>                                                                                                        |      |
| Mo Zhiqiang, Yang Hua, Jiahui Li, Yuan Liu, Shawn Wong and Jianmin Huang .....                                                                                                                             | 5304 |
| <i>Add-One-In: Incremental Sample Selection for Large Language Models via a Choice-Based Greedy Paradigm</i>                                                                                               |      |
| Zhuo Li, Yuhao Du, Xiaoqi Jiao, Steven Y. Guo, Yuege Feng, Xiang Wan, Anningzhe Gao and Jinpeng Hu .....                                                                                                   | 5322 |
| <i>QuZO: Quantized Zeroth-Order Fine-Tuning for Large Language Models</i>                                                                                                                                  |      |
| Jiajun Zhou, Yifan Yang, Kai Zhen, Ziyue Liu, Yequan Zhao, Ershad Banijamali, Athanasios Mouchtaris, Ngai Wong and Zheng Zhang .....                                                                       | 5342 |
| <i>Cost-Optimal Grouped-Query Attention for Long-Context Modeling</i>                                                                                                                                      |      |
| Yingfa Chen, Yutong Wu, Chenyang Song, Zhen Leng Thai, Xingyu Shen, Xu Han, Zhiyuan Liu and Maosong Sun .....                                                                                              | 5361 |
| <i>ChatVLA: Unified Multimodal Understanding and Robot Control with Vision-Language-Action Model</i>                                                                                                       |      |
| Zhongyi Zhou, Yichen Zhu, Minjie Zhu, Junjie Wen, Ning Liu, Zhiyuan Xu, Weibin Meng, Yaxin Peng, Chaomin Shen, Feifei Feng and Yi Xu .....                                                                 | 5378 |
| <i>KG-RAG: Enhancing GUI Agent Decision-Making via Knowledge Graph-Driven Retrieval-Augmented Generation</i>                                                                                               |      |
| Ziyi Guan, Jason Chun Lok Li, Zhijian Hou, Pingping Zhang, Donglai Xu, Yuzhi Zhao, Mengyang Wu, Jinpeng Chen, Thanh-Toan Nguyen, Pengfei Xian, Wenao Ma, Shengchao Qin, Graziano Chesi and Ngai Wong ..... | 5397 |
| <i>CLIP-MoE: Towards Building Mixture of Experts for CLIP with Diversified Multiplier Upcycling</i>                                                                                                        |      |
| Jihai Zhang, Xiaoye Qu, Tong Zhu and Yu Cheng .....                                                                                                                                                        | 5407 |
| <i>Search-o1: Agentic Search-Enhanced Large Reasoning Models</i>                                                                                                                                           |      |
| Xiaoxi Li, Guanting Dong, Jiajie Jin, Yuyao Zhang, Yujia Zhou, Yutao Zhu, Peitian Zhang and Zhicheng Dou .....                                                                                             | 5421 |
| <i>From Personas to Talks: Revisiting the Impact of Personas on LLM-Synthesized Emotional Support Conversations</i>                                                                                        |      |
| Shenghan Wu, Yimo Zhu, Wynne Hsu, Mong-Li Lee and Yang Deng .....                                                                                                                                          | 5440 |
| <i>Select-Then-Decompose: From Empirical Analysis to Adaptive Selection Strategy for Task Decomposition in Large Language Models</i>                                                                       |      |
| Shuodi Liu, Yingzhuo Liu, Zi Wang, Yusheng Wang, Huijia Wu, Liuyu Xiang and Zhaofeng He .....                                                                                                              | 5455 |
| <i>TombRaider: Entering the Vault of History to Jailbreak Large Language Models</i>                                                                                                                        |      |
| Junchen Ding, Jiahao Zhang, Yi Liu, Ziqi Ding, Gelei Deng and Yuekang Li .....                                                                                                                             | 5479 |
| <i>Text Meets Topology: Rethinking Out-of-distribution Detection in Text-Rich Networks</i>                                                                                                                 |      |
| Danny Wang, Ruihong Qiu, Guangdong Bai and Zi Huang .....                                                                                                                                                  | 5495 |
| <i>APLOT: Robust Reward Modeling via Adaptive Preference Learning with Optimal Transport</i>                                                                                                               |      |
| Zhuo Li, Yuege Feng, Dandan Guo, Jinpeng Hu, Anningzhe Gao and Xiang Wan .....                                                                                                                             | 5525 |

|                                                                                                                                                                           |      |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>HS-STaR: Hierarchical Sampling for Self-Taught Reasoners via Difficulty Estimation and Budget Reallocation</i>                                                         |      |
| Feng Xiong, Hongling Xu, Yifei Wang, Runxi Cheng, Yong Wang and Xiangxiang Chu . . .                                                                                      | 5540 |
| <i>SEPS: A Separability Measure for Robust Unlearning in LLMs</i>                                                                                                         |      |
| Wonje Jeung, Sangyeon Yoon and Albert No . . . . .                                                                                                                        | 5557 |
| <i>TRUST-VL: An Explainable News Assistant for General Multimodal Misinformation Detection</i>                                                                            |      |
| Zehong Yan, Peng Qi, Wynne Hsu and Mong-Li Lee . . . . .                                                                                                                  | 5589 |
| <i>Tree-of-Quote Prompting Improves Factuality and Attribution in Multi-Hop and Medical Reasoning</i>                                                                     |      |
| Justin Xu, Yiming Li, Zizheng Zhang, Augustine Yui Hei Luk, Mayank Jobanputra, Samarth Oza, Ashley Murray, Meghana Reddy Kasula, Andrew Parker and David W Eyre . . . . . | 5606 |
| <i>UnitCoder: Scalable Code Synthesis from Pre-training Corpora</i>                                                                                                       |      |
| Yichuan Ma, Yunfan Shao, Peiji Li, Demin Song, Qipeng Guo, Linyang Li, Xipeng Qiu and Kai Chen . . . . .                                                                  | 5624 |
| <i>GRPO-LEAD: A Difficulty-Aware Reinforcement Learning Approach for Concise Mathematical Reasoning in Language Models</i>                                                |      |
| Jixiao Zhang and Chunsheng Zuo . . . . .                                                                                                                                  | 5643 |
| <i>Improving Low-Resource Sequence Labeling with Knowledge Fusion and Contextual Label Explanations</i>                                                                   |      |
| Peichao Lai, Jiaxin Gan, Feiyang Ye, Wentao Zhang, Fangcheng Fu, Yilei Wang and Bin Cui                                                                                   | 5667 |
| <i>Rethinking Cross-Subject Data Splitting for Brain-to-Text Decoding</i>                                                                                                 |      |
| Congchi Yin, Qian Yu, Zhiwei Fang, Changping Peng and Piji Li . . . . .                                                                                                   | 5687 |
| <i>RCScore: Quantifying Response Consistency in Large Language Models</i>                                                                                                 |      |
| Dongjun Jang, Youngchae Ahn and Hyopil Shin . . . . .                                                                                                                     | 5702 |
| <i>A Multi-Agent Framework with Automated Decision Rule Optimization for Cross-Domain Misinformation Detection</i>                                                        |      |
| Hui Li, Ante Wang, Kunquan Li, Zhihao Wang, Liang Zhang, Delai Qiu, Qingsong Liu and Jinsong Su . . . . .                                                                 | 5721 |
| <i>OmniEval: An Omnidirectional and Automatic RAG Evaluation Benchmark in Financial Domain</i>                                                                            |      |
| Shuting Wang, Jiejun Tan, Zhicheng Dou and Ji-Rong Wen . . . . .                                                                                                          | 5738 |
| <i>AQuilt: Weaving Logic and Self-Inspection into Low-Cost, High-Relevance Data Synthesis for Specialist LLMs</i>                                                         |      |
| Xiaopeng Ke, Hexuan Deng, Xuebo Liu, Jun Rao, Zhenxi Song, Jun Yu and Min Zhang . . .                                                                                     | 5764 |
| <i>MoSEs: Uncertainty-Aware AI-Generated Text Detection via Mixture of Stylistics Experts with Conditional Thresholds</i>                                                 |      |
| Junxi Wu, Jinpeng Wang, Zheng Liu, Bin Chen, Dongjian Hu, Hao Wu and Shu-Tao Xia . . .                                                                                    | 5798 |
| <i>Merger-as-a-Stealer: Stealing Targeted PII from Aligned LLMs with Model Merging</i>                                                                                    |      |
| Lin Lu, Zhigang Zuo, Ziji Sheng and Pan Zhou . . . . .                                                                                                                    | 5818 |
| <i>Pragmatic Inference Chain (PIC) Improving LLMs' Reasoning of Authentic Implicit Toxic Language</i>                                                                     |      |
| Xi Chen and Shuo Wang . . . . .                                                                                                                                           | 5838 |
| <i>Beyond Demonstrations: Dynamic Vector Construction from Latent Representations</i>                                                                                     |      |
| Wang Cai, Hsiu-Yuan Huang, Zhixiang Wang and Yunfang Wu . . . . .                                                                                                         | 5854 |

|                                                                                                                                                                                                                      |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Detoxifying Large Language Models via the Diversity of Toxic Samples</i>                                                                                                                                          |      |
| Ying Zhao, Yuanzhao Guo, Xuemeng Weng, Yuan Tian, Wei Wang and Yi Chang . . . . .                                                                                                                                    | 5870 |
| <i>LLM-Driven Implicit Target Augmentation and Fine-Grained Contextual Modeling for Zero-Shot and Few-Shot Stance Detection</i>                                                                                      |      |
| Yanxu Ji, Jinzhong Ning, Yijia Zhang, Zhi Liu and Hongfei Lin . . . . .                                                                                                                                              | 5884 |
| <i>Dial-In LLM: Human-Aligned LLM-in-the-loop Intent Clustering for Customer Service Dialogues</i>                                                                                                                   |      |
| Mengze Hong, Wailing Ng, Chen Jason Zhang, Yuanfeng Song and Di Jiang . . . . .                                                                                                                                      | 5897 |
| <i>Superficial Self-Improved Reasoners Benefit from Model Merging</i>                                                                                                                                                |      |
| Xiangchi Yuan, Chunhui Zhang, Zheyuan Liu, Dachuan Shi, Leyan Pan, Soroush Vosoughi and Wenke Lee . . . . .                                                                                                          | 5913 |
| <i>CARFT: Boosting LLM Reasoning via Contrastive Learning with Annotated Chain-of-Thought-based Reinforced Fine-Tuning</i>                                                                                           |      |
| Wenqiao Zhu, Ji Liu, Rongjunchen Zhang, Haipang WU and Yulun Zhang . . . . .                                                                                                                                         | 5934 |
| <i>QualBench: Benchmarking Chinese LLMs with Localized Professional Qualifications for Vertical Domain Evaluation</i>                                                                                                |      |
| Mengze Hong, Wailing Ng, Chen Jason Zhang and Di Jiang . . . . .                                                                                                                                                     | 5950 |
| <i>VideoEraser: Concept Erasure in Text-to-Video Diffusion Models</i>                                                                                                                                                |      |
| Naen Xu, Jinghuai Zhang, Changjiang Li, Zhi Chen, Chunyi Zhou, Qingming Li, Tianyu Du and Shouling Ji . . . . .                                                                                                      | 5966 |
| <i>Diagram-Driven Course Questions Generation</i>                                                                                                                                                                    |      |
| Xinyu Zhang, Lingling Zhang, Yanrui Wu, Muye Huang, Wenjun Wu, Bo Li, Shaowei Wang, Basura Fernando and Jun Liu . . . . .                                                                                            | 5996 |
| <i>ECC: An Emotion-Cause Conversation Dataset for Empathy Response</i>                                                                                                                                               |      |
| Yuanyuan He, Yongsan Pan, Wei Li, Jiali You, Jiawen Deng and Fuji Ren . . . . .                                                                                                                                      | 6012 |
| <i>ThoughtProbe: Classifier-Guided LLM Thought Space Exploration via Probing Representations</i>                                                                                                                     |      |
| Zijian Wang and Chang Xu . . . . .                                                                                                                                                                                   | 6030 |
| <i>JOLT-SQL: Joint Loss Tuning of Text-to-SQL with Confusion-aware Noisy Schema Sampling</i>                                                                                                                         |      |
| Jinwang Song, Hongying Zan, Kunli Zhang, Lingling Mu, Yingjie Han, Haobo Hua and Min Peng . . . . .                                                                                                                  | 6052 |
| <i>DMDTEval: An Evaluation and Analysis of LLMs on Disambiguation in Multi-domain Translation</i>                                                                                                                    |      |
| Zhibo Man, Yuanmeng Chen, Yujie Zhang and Jinan Xu . . . . .                                                                                                                                                         | 6066 |
| <i>SciRIF: A Resource to Enhance Language Model Instruction-Following over Scientific Literature</i>                                                                                                                 |      |
| David Wadden, Kejian Shi, Jacob Morrison, Alan Li, Aakanksha Naik, Shruti Singh, Nitzan Barzilay, Kyle Lo, Tom Hope, Luca Soldaini, Shannon Zejiang Shen, Doug Downey, Hannaneh Hajishirzi and Arman Cohan . . . . . | 6084 |
| <i>MAKAR: a Multi-Agent framework based Knowledge-Augmented Reasoning for Grounded Multimodal Named Entity Recognition</i>                                                                                           |      |
| Xinkui Lin, Yuhui Zhang, Yongxiu Xu, Kun Huang, Hongzhang Mu, Yubin Wang, Gaopeng Gou, Li Qian, Li Peng, Wei Liu, Jian Luan and Hongbo Xu . . . . .                                                                  | 6122 |
| <i>VisCRA: A Visual Chain Reasoning Attack for Jailbreaking Multimodal Large Language Models</i>                                                                                                                     |      |
| Bingrui Sima, Linhua Cong, Wenxuan Wang and Kun He . . . . .                                                                                                                                                         | 6143 |



|                                                                                                                                                                                                      |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Investigating Neurons and Heads in Transformer-based LLMs for Typographical Errors</i>                                                                                                            |      |
| Kohei Tsuji, Tatsuya Hiraoka, Yuchang Cheng, Eiji Aramaki and Tomoya Iwakura . . . . .                                                                                                               | 6157 |
| <i>LMR-BENCH: Evaluating LLM Agent’s Ability on Reproducing Language Modeling Research</i>                                                                                                           |      |
| Shuo Yan, Ruochen Li, Ziming Luo, Zimu Wang, Daoyang Li, Liqiang Jing, Kaiyu He, Peilin Wu, Juntong Ni, George Michalopoulos, Yue Zhang, Ziyang Zhang, Mian Zhang, Zhiyu Chen and Xinya Du . . . . . | 6176 |
| <i>RAV: Retrieval-Augmented Voting for Tactile Descriptions Without Training</i>                                                                                                                     |      |
| Jinlin Wang, Yulong Ji and Hongyu Yang . . . . .                                                                                                                                                     | 6199 |
| <i>Static Word Embeddings for Sentence Semantic Representation</i>                                                                                                                                   |      |
| Takashi Wada, Yuki Hirakawa, Ryotaro Shimizu, Takahiro Kawashima and Yuki Saito . . . . .                                                                                                            | 6207 |
| <i>PropRAG: Guiding Retrieval with Beam Search over Proposition Paths</i>                                                                                                                            |      |
| Jingjin Wang and Jiawei Han . . . . .                                                                                                                                                                | 6224 |
| <i>Rethinking Backdoor Detection Evaluation for Language Models</i>                                                                                                                                  |      |
| Jun Yan, Wenjie Jacky Mo, Xiang Ren and Robin Jia . . . . .                                                                                                                                          | 6240 |
| <i>Glider: Global and Local Instruction-Driven Expert Router</i>                                                                                                                                     |      |
| Pingzhi Li, Prateek Yadav, Jaehong Yoon, Jie Peng, Yi-Lin Sung, Mohit Bansal and Tianlong Chen . . . . .                                                                                             | 6252 |
| <i>CoVoGER: A Multilingual Multitask Benchmark for Speech-to-text Generative Error Correction with Large Language Models</i>                                                                         |      |
| Zhengdong Yang, Zhen Wan, Sheng Li, Chao-Han Huck Yang and Chenhui Chu . . . . .                                                                                                                     | 6314 |
| <i>Tiny Budgets, Big Gains: Parameter Placement Strategy in Parameter Super-Efficient Fine-Tuning</i>                                                                                                |      |
| Jinman Zhao, Xueyan Zhang, Jiaru Li, Jingcheng Niu, Yulan Hu, Erxue Min and Gerald Penn                                                                                                              | 6327 |
| <i>Legal Fact Prediction: The Missing Piece in Legal Judgment Prediction</i>                                                                                                                         |      |
| Junkai Liu, Yujie Tong, Hui Huang, Bowen Zheng, Yiran HU, Peicheng Wu, Chuan Xiao, Makoto Onizuka, Muyun Yang and Shuyuan Zheng . . . . .                                                            | 6346 |
| <i>DAMON: A Dialogue-Aware MCTS Framework for Jailbreaking Large Language Models</i>                                                                                                                 |      |
| Xu Zhang, Xunjian Yin, Dinghao Jing, Huixuan Zhang, Xinyu Hu and Xiaojun Wan . . . . .                                                                                                               | 6362 |
| <i>Multilingual Prompting for Improving LLM Generation Diversity</i>                                                                                                                                 |      |
| Qihan Wang, Shidong Pan, Tal Linzen and Emily Black . . . . .                                                                                                                                        | 6379 |
| <i>MOSAIC: Modeling Social AI for Content Dissemination and Regulation in Multi-Agent Simulations</i>                                                                                                |      |
| Genglin Liu, Vivian T. Le, Salman Rahman, Elisa Kreiss, Marzyeh Ghassemi and Saadia Gabriel                                                                                                          | 6402 |
| <i>Identification of Multiple Logical Interpretations in Counter-Arguments</i>                                                                                                                       |      |
| Wenzhi Wang, Paul Reisert, Shoichi Naito, Naoya Inoue, Machi Shimmei, Surawat Pothong, Jungmin Choi and Kentaro Inui . . . . .                                                                       | 6430 |
| <i>LyapLock: Bounded Knowledge Preservation in Sequential Large Language Model Editing</i>                                                                                                           |      |
| Peng Wang, Biyu Zhou, Xuehai Tang, Jizhong Han and Songlin Hu . . . . .                                                                                                                              | 6446 |
| <i>AlignX: Advancing Multilingual Large Language Models with Multilingual Representation Alignment</i>                                                                                               |      |
| Mengyu Bu, Shaolei Zhang, Zhongjun He, Hua Wu and Yang Feng . . . . .                                                                                                                                | 6472 |

|                                                                                                                                                    |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>What Makes a Good Reasoning Chain? Uncovering Structural Patterns in Long Chain-of-Thought Reasoning</i>                                        |      |
| Gangwei Jiang, Yahui Liu, Zhaoyi Li, Wei Bi, Fuzheng Zhang, Linqi Song, Ying Wei and Defu Lian .....                                               | 6502 |
| <i>HD-PiSSA: High-Rank Distributed Orthogonal Adaptation</i>                                                                                       |      |
| Yiding Wang, Fanxu Meng, Xuefeng Zhang, Fan Jiang, Pingzhi Tang and Muhan Zhang ...                                                                | 6527 |
| <i>Firewall Routing: Blocking Leads to Better Hybrid Inference for LLMs</i>                                                                        |      |
| Runyu Peng, Yunhua Zhou, Kai Lv, Yang Gao, Qipeng Guo and Xipeng Qiu .....                                                                         | 6541 |
| <i>SPE Attention: Making Attention Equivariant to Semantic-Preserving Permutation for Code Processing</i>                                          |      |
| Chengyu Jiao, Shuhao Chen and Yu Zhang .....                                                                                                       | 6567 |
| <i>Audio-centric Video Understanding Benchmark without Text Shortcut</i>                                                                           |      |
| Yudong Yang, Jimin Zhuang, Guangzhi Sun, Changli Tang, Yixuan Li, Peihan Li, Yifan Jiang, Wei Li, Zejun MA and Chao Zhang .....                    | 6581 |
| <i>TurboRAG: Accelerating Retrieval-Augmented Generation with Precomputed KV Caches for Chunked Text</i>                                           |      |
| Songshuo Lu, Hua Wang, Yutian Rong, Zhi Chen and Yaohua Tang .....                                                                                 | 6600 |
| <i>ZoomEye: Enhancing Multimodal LLMs with Human-Like Zooming Capabilities through Tree-Based Image Exploration</i>                                |      |
| Haozhan Shen, Kangjia Zhao, Tiancheng Zhao, Ruochen Xu, Zilun Zhang, Mingwei Zhu and Jianwei Yin .....                                             | 6614 |
| <i>Learning Like Humans: Advancing LLM Reasoning Capabilities via Adaptive Difficulty Curriculum Learning and Expert-Guided Self-Reformulation</i> |      |
| Enci Zhang, Xingang Yan, Wei Lin, Tianxiang. Zhang and LU Qianchun .....                                                                           | 6631 |
| <i>VersaTune: An Efficient Data Composition Framework for Training Multi-Capability LLMs</i>                                                       |      |
| Keer Lu, Keshi Zhao, Zhuoran Zhang, Zheng Liang, Bin Cui, Tengjiao Wang and Wentao Zhang                                                           | 6646 |
| <i>FlightGPT: Towards Generalizable and Interpretable UAV Vision-and-Language Navigation with Vision-Language Models</i>                           |      |
| Hengxing Cai, Jinhan Dong, Jingjun Tan, Jingcheng Deng, Sihang Li, Zhifeng Gao, Haidong Wang, Zicheng Su, Agachai Sumalee and Renxin Zhong .....   | 6671 |
| <i>Multimodal Language Models See Better When They Look Shallower</i>                                                                              |      |
| Haoran Chen, Junyan Lin, Xinghao Chen, Yue Fan, Jianfeng Dong, Xin Jin, Hui Su, Jinlan Fu and Xiaoyu Shen .....                                    | 6689 |
| <i>LoSiA: Efficient High-Rank Fine-Tuning via Subnet Localization and Optimization</i>                                                             |      |
| Xujia Wang, Yunjia Qi and Bin Xu .....                                                                                                             | 6708 |
| <i>Invisible Entropy: Towards Safe and Efficient Low-Entropy LLM Watermarking</i>                                                                  |      |
| Tianle Gu, Zongqi Wang, Kexin Huang, Yuanqi Yao, Xiangliang Zhang, Yujiu Yang and Xiuying Chen .....                                               | 6728 |
| <i>Measuring Bias or Measuring the Task: Understanding the Brittle Nature of LLM Gender Biases</i>                                                 |      |
| Bufan Gao and Elisa Kreiss .....                                                                                                                   | 6746 |
| <i>Alignment-Augmented Speculative Decoding with Alignment Sampling and Conditional Verification</i>                                               |      |
| Jikai Wang, Zhenxu Tian, Juntao Li, Qingrong Xia, Xinyu Duan, Zhefeng Wang, Baoxing Huai and Min Zhang .....                                       | 6763 |



|                                                                                                                                                                                           |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>ViLBench: A Suite for Vision-Language Process Reward Modeling</i>                                                                                                                      |      |
| Haoqin Tu, Weitao Feng, Hardy Chen, Hui Liu, Xianfeng Tang and Cihang Xie . . . . .                                                                                                       | 6776 |
| <i>Keep Security! Benchmarking Security Policy Preservation in Large Language Model Contexts Against Indirect Attacks in Question Answering</i>                                           |      |
| Hwan Chang, Yumin Kim, Yonghyun Jun and Hwanhee Lee . . . . .                                                                                                                             | 6792 |
| <i>Route Sparse Autoencoder to Interpret Large Language Models</i>                                                                                                                        |      |
| Wei Shi, Sihang Li, Tao Liang, Mingyang Wan, Guojun Ma, Xiang Wang and Xiangnan He                                                                                                        | 6813 |
| <i>BTS: Harmonizing Specialized Experts into a Generalist LLM</i>                                                                                                                         |      |
| Qizhen Zhang, Prajjwal Bhargava, Chloe Bi, Chris X. Cai, Jakob Nicolaus Foerster, Jeremy Fu, Punit Singh Koura, Ruan Silva, Sheng Shen, Emily Dinan, Suchin Gururangan and Mike Lewis . . | 6828 |
| <i>CoCoA: Confidence- and Context-Aware Adaptive Decoding for Resolving Knowledge Conflicts in Large Language Models</i>                                                                  |      |
| Anant Khandelwal, Manish Gupta and Puneet Agrawal . . . . .                                                                                                                               | 6847 |
| <i>R-Bind: Unified Enhancement of Attribute and Relation Binding in Text-to-Image Diffusion Models</i>                                                                                    |      |
| Huixuan Zhang and Xiaojun Wan . . . . .                                                                                                                                                   | 6868 |
| <i>Middo: Model-Informed Dynamic Data Optimization for Enhanced LLM Fine-Tuning via Closed-Loop Learning</i>                                                                              |      |
| Zinan Tang, Xin Gao, Qizhi Pei, Zhuoshi Pan, Mengzhang Cai, Jiang Wu, Conghui He and Lijun Wu . . . . .                                                                                   | 6883 |
| <i>Information Integration in Large Language Models is Gated by Linguistic Structural Markers</i>                                                                                         |      |
| Wei Liu and Nai Ding . . . . .                                                                                                                                                            | 6904 |
| <i>Why and How LLMs Benefit from Knowledge Introspection in Commonsense Reasoning</i>                                                                                                     |      |
| Chengfeng Zhao, Shizhu He, Shanshan Jiang, Bin Dong, Jun Zhao and Kang Liu . . . . .                                                                                                      | 6917 |
| <i>GraDaSE: Graph-Based Dataset Search with Examples</i>                                                                                                                                  |      |
| Jing He, Mingyang Lv, Qing Shi and Gong Cheng . . . . .                                                                                                                                   | 6933 |
| <i>Confidence-guided Refinement Reasoning for Zero-shot Question Answering</i>                                                                                                            |      |
| Youwon Jang, Woo Suk Choi, Minjoon Jung, Minsu Lee and Byoung-Tak Zhang . . . . .                                                                                                         | 6945 |
| <i>DICE: Structured Reasoning in LLMs through SLM-Guided Chain-of-Thought Correction</i>                                                                                                  |      |
| Yiqi Li, Yusheng Liao, Zhe Chen, Yanfeng Wang and Yu Wang . . . . .                                                                                                                       | 6963 |
| <i>CTCC: A Robust and Stealthy Fingerprinting Framework for Large Language Models via Cross-Turn Contextual Correlation Backdoor</i>                                                      |      |
| Zhenhua Xu, Xixiang Zhao, Xubin Yue, Shengwei Tian, Changting Lin and Meng Han . . . .                                                                                                    | 6979 |
| <i>Realistic Training Data Generation and Rule Enhanced Decoding in LLM for NameGuess</i>                                                                                                 |      |
| Yikuan Xia, Jiazun Chen, Sujian Li and Jun Gao . . . . .                                                                                                                                  | 7002 |
| <i>EverTracer: Hunting Stolen Large Language Models via Stealthy and Robust Probabilistic Fingerprint</i>                                                                                 |      |
| Zhenhua Xu, Meng Han and Wenpeng Xing . . . . .                                                                                                                                           | 7020 |
| <i>Selective Preference Optimization via Token-Level Reward Function Estimation</i>                                                                                                       |      |
| Kailai Yang, Zhiwei Liu, Qianqian Xie, Jimin Huang, Erxue Min and Sophia Ananiadou . .                                                                                                    | 7044 |
| <i>Arena-lite: Efficient and Reliable Large Language Model Evaluation via Tournament-Based Direct Comparisons</i>                                                                         |      |
| Seonil Son, Ju-Min Oh, Heegon Jin, Cheolhun Jang, Jeongbeom Jeong and KunTae Kim . .                                                                                                      | 7069 |

|                                                                                                                                         |      |
|-----------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Addressing Tokenization Inconsistency in Steganography and Watermarking Based on Large Language Models</i>                           |      |
| Ruiyi Yan and Yugo Murawaki .....                                                                                                       | 7088 |
| <i>ExeCoder: Empowering Large Language Models with Executability Representation for Code Translation</i>                                |      |
| Minghua He, Yue Chen, Fangkai Yang, Pu Zhao, Wenjie Yin, Yu Kang, Qingwei Lin, Saravan Rajmohan and Dongmei Zhang .....                 | 7111 |
| <i>TableEval: A Real-World Benchmark for Complex, Multilingual, and Multi-Structured Table Question Answering</i>                       |      |
| Junnan Zhu, Jingyi Wang, Bohan Yu, Xiaoyu Wu, Junbo Li, Lei Wang and Nan Xu .....                                                       | 7138 |
| <i>NOVA-63: Native Omni-lingual Versatile Assessments of 63 Disciplines</i>                                                             |      |
| Jinyang Zhang, Kexin Yang, Yu Wan, Muyang Ye, Baosong Yang, Fei Huang, Junyang Lin and Dayiheng Liu .....                               | 7159 |
| <i>InfoGain-RAG: Boosting Retrieval-Augmented Generation through Document Information Gain-based Reranking and Filtering</i>            |      |
| Zihan Wang, Zihan Liang, Zhou Shao, Yufei Ma, Huangyu Dai, Ben Chen, Lingtao Mao, Chenyi Lei, Yuqing Ding and Han Li .....              | 7202 |
| <i>SpecVLM: Enhancing Speculative Decoding of Video LLMs via Verifier-Guided Token Pruning</i>                                          |      |
| Yicheng Ji, Jun Zhang, Heming Xia, Jinpeng Chen, Lidan Shou, Gang Chen and Huan Li ..                                                   | 7217 |
| <i>What Do Indonesians Really Need from Language Technology? A Nationwide Survey</i>                                                    |      |
| Muhammad Dehan Al Kautsar, Lucky Susanto, Derry Tanti Wijaya and Fajri Koto .....                                                       | 7232 |
| <i>LEO-MINI: An Efficient Multimodal Large Language Model using Conditional Token Reduction and Mixture of Multi-Modal Experts</i>      |      |
| Yimu Wang, Mozghan Nasr Azadani, Sean Sedwards and Krzysztof Czarnecki .....                                                            | 7258 |
| <i>Confounding Factors in Relating Model Performance to Morphology</i>                                                                  |      |
| Wessel Poelman, Thomas Bauwens and Miryam de Lhoneux .....                                                                              | 7274 |
| <i>Context-Aware Membership Inference Attacks against Pre-trained Large Language Models</i>                                             |      |
| Hongyan Chang, Ali Shahin Shamsabadi, Kleomenis Katevas, Hamed Haddadi and Reza Shokri                                                  | 7300 |
| <i>Formalizing Style in Personal Narratives</i>                                                                                         |      |
| Gustave Cortal and Alain Finkel .....                                                                                                   | 7323 |
| <i>TopicAttack: An Indirect Prompt Injection Attack via Topic Transition</i>                                                            |      |
| Yulin Chen, Haoran Li, Yuexin Li, Yue Liu, Yangqiu Song and Bryan Hooi .....                                                            | 7339 |
| <i>PSET: a Phonetics-Semantics Evaluation Testbed</i>                                                                                   |      |
| Gianluca Sperduti and Dong Nguyen .....                                                                                                 | 7358 |
| <i>From Unaligned to Aligned: Scaling Multilingual LLMs with Multi-Way Parallel Corpora</i>                                             |      |
| Yingli Shen, Wen Lai, Shuo Wang, Ge Gao, Kangyang Luo, Alexander Fraser and Maosong Sun                                                 | 7369 |
| <i>GATEAU: Selecting Influential Samples for Long Context Alignment</i>                                                                 |      |
| Shuzheng Si, Haozhe Zhao, Gang Chen, Yunshui Li, Kangyang Luo, Chuancheng Lv, Kaikai An, Fanchao Qi, Baobao Chang and Maosong Sun ..... | 7392 |

|                                                                                                                                          |      |
|------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Teach Small Models to Reason by Curriculum Distillation</i>                                                                           |      |
| Wangyi Jiang, Yaojie Lu, Hongyu Lin, Xianpei Han and Le Sun .....                                                                        | 7424 |
| <i>Enhancing Reasoning Abilities of Small LLMs with Cognitive Alignment</i>                                                              |      |
| Wenrui Cai, Chengyu Wang, Junbing Yan, Jun Huang and Xiangzhong Fang .....                                                               | 7435 |
| <i>NOVER: Incentive Training for Language Models via Verifier-Free Reinforcement Learning</i>                                            |      |
| Wei Liu, Siya Qi, Xinyu Wang, Chen Qian, Yali Du and Yulan He .....                                                                      | 7451 |
| <i>Genre Matters: How Text Types Interact with Decoding Strategies and Lexical Predictors in Shaping Reading Behavior</i>                |      |
| Lena Sophia Bolliger and Lena Ann Jäger .....                                                                                            | 7471 |
| <i>RTE-GMoE: A Model-agnostic Approach for Relation Triplet Extraction via Graph-based Mixture-of-Expert Mutual Learning</i>             |      |
| Aziguli Wulamu, Kaiyuan Gong, Lyu Zhengyu, Yu Han, Zhihong Zhu and Bowen Xing ...                                                        | 7489 |
| <i>Avoidance Decoding for Diverse Multi-Branch Story Generation</i>                                                                      |      |
| Kyeongman Park, Nakyeong Yang and Kyomin Jung .....                                                                                      | 7501 |
| <i>Probabilistic Soundness Guarantees in LLM Reasoning Chains</i>                                                                        |      |
| Weiqiu You, Anton Xue, Shreya Havaldar, Delip Rao, Helen Jin, Chris Callison-Burch and Eric Wong .....                                   | 7518 |
| <i>SQLWOZ: A Realistic Task-Oriented Dialogue Dataset with SQL-Based Dialogue State Representation for Complex User Requirements</i>     |      |
| Heng-Da Xu, Xian-Ling Mao, Fanshu Sun, Tian-Yi Che, Cheng-Xin Xin and Heyan Huang                                                        | 7538 |
| <i>SURE: Safety Understanding and Reasoning Enhancement for Multimodal Large Language Models</i>                                         |      |
| Yuxin Gou, Xiaoning Dong, Qin Li, Shishen Gu, Richang Hong and Wenbo Hu .....                                                            | 7564 |
| <i>EMO: Embedding Model Distillation via Intra-Model Relation and Optimal Transport Alignments</i>                                       |      |
| Minh-Phuc Truong, Hai An Vu, Tu Vu, Nguyen Thi Ngoc Diep, Linh Ngo Van, Thien Huu Nguyen and Trung Le .....                              | 7606 |
| <i>AesBiasBench: Evaluating Bias and Alignment in Multimodal Language Models for Personalized Image Aesthetic Assessment</i>             |      |
| Kun Li, Lai Man Po, Hongzheng Yang, Xuyuan XU, Kangcheng Liu and Yuzhi Zhao .....                                                        | 7619 |
| <i>DA-Pred: Performance Prediction for Text Summarization under Domain-Shift and Instruct-Tuning</i>                                     |      |
| Anum Afzal, Florian Matthes and Alexander Fabbri .....                                                                                   | 7633 |
| <i>UnCo: Uncertainty-Driven Collaborative Framework of Large and Small Models for Grounded Multimodal NER</i>                            |      |
| Jielong Tang, Yang Yang, Jianxing Yu, Zhen-Xing Wang, Haoyuan Liang, Liang Yao and Jian Yin .....                                        | 7645 |
| <i>An Empirical Study of LLM Reasoning Ability Under Strict Output Length Constraint</i>                                                 |      |
| Yi Sun, Han Wang, Jiaqiang Li, Jiacheng Liu, Xiangyu Li, Hao Wen, Yizhen Yuan, Huiwen Zheng, Yan Liang, Yuanchun Li and Yunxin Liu ..... | 7664 |
| <i>Enrich-on-Graph: Query-Graph Alignment for Complex Reasoning with LLM Enriching</i>                                                   |      |
| Songze Li, Zhiqiang Liu, Zhengke Gui, Huajun Chen and Wen Zhang .....                                                                    | 7684 |
| <i>Noise, Adaptation, and Strategy: Assessing LLM Fidelity in Decision-Making</i>                                                        |      |
| Yuanjun Feng, Vivek Choudhary and Yash Raj Shrestha .....                                                                                | 7705 |

|                                                                                                                                                     |      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>Structuring Radiology Reports: Challenging LLMs with Lightweight Models</i>                                                                      |      |
| Johannes Moll, Louisa Fay, Asfandiyar Azhar, Sophie Ostmeier, Sergios Gatidis, Tim C. Lueth, Curtis Langlotz and Jean-Benoit Delbrouck . . . . .    | 7719 |
| <i>PricingLogic: Evaluating LLMs Reasoning on Complex Tourism Pricing Tasks</i>                                                                     |      |
| Yunuo Liu, Dawei Zhu, Zena Al-Khalili, Dai Cheng, Yanjun Chen, Dietrich Klakow, Wei Zhang and Xiaoyu Shen . . . . .                                 | 7737 |
| <i>EcoTune: Token-Efficient Multi-Fidelity Hyperparameter Optimization for Large Language Model Inference</i>                                       |      |
| Yuebin XU, Zhiyi Chen and Zeyi Wen . . . . .                                                                                                        | 7747 |
| <i>Investigating Value-Reasoning Reliability in Small Large Language Models</i>                                                                     |      |
| Xia Du, Shuhan Sun, Pengyuan Liu and Dong Yu . . . . .                                                                                              | 7758 |
| <i>Can LLMs Explain Themselves Counterfactually?</i>                                                                                                |      |
| Zahra Dehghanighobadi, Asja Fischer and Muhammad Bilal Zafar . . . . .                                                                              | 7799 |
| <i>Self-Adjust Softmax</i>                                                                                                                          |      |
| Chuanyang Zheng, Yihang Gao, Guoxuan Chen, Han Shi, Jing Xiong, Xiaozhe Ren, Chao Huang, Zhenguo Li and Yu Li . . . . .                             | 7828 |
| <i>DiscoSG: Towards Discourse-Level Text Scene Graph Parsing through Iterative Graph Refinement</i>                                                 |      |
| Shaoqing Lin, Chong Teng, Fei Li, Donghong Ji, Lizhen Qu and Zhuang Li . . . . .                                                                    | 7849 |
| <i>XAutoLM: Efficient Fine-Tuning of Language Models via Meta-Learning and AutoML</i>                                                               |      |
| Ernesto Luis Estevanell Valladares, Suilan Estevez-Velarde, Yoan Gutierrez, Andrés Montoyo and Ruslan Mitkov . . . . .                              | 7875 |
| <i>UNCERTAINTY-LINE: Length-Invariant Estimation of Uncertainty for Large Language Models</i>                                                       |      |
| Roman Vashurin, Maiya Goloburda, Preslav Nakov and Maxim Panov . . . . .                                                                            | 7893 |
| <i>WebAgent-R1: Training Web Agents via End-to-End Multi-Turn Reinforcement Learning</i>                                                            |      |
| Zhepei Wei, Wenlin Yao, Yao Liu, Weizhi Zhang, Qin Lu, Liang Qiu, Changlong Yu, Puyang Xu, Chao Zhang, Bing Yin, Hyokun Yun and Lihong Li . . . . . | 7921 |
| <i>Same evaluation, more tokens: On the effect of input length for machine translation evaluation using Large Language Models</i>                   |      |
| Tobias Domhan and Dawei Zhu . . . . .                                                                                                               | 7941 |
| <i>PAKTON: A Multi-Agent Framework for Question Answering in Long Legal Agreements</i>                                                              |      |
| Raptopoulos Petros, Giorgos Filandrianos, Maria Lymperaïou and Giorgos Stamou . . . . .                                                             | 7960 |
| <i>PoSum-Bench: Benchmarking Position Bias in LLM-based Conversational Summarization</i>                                                            |      |
| XU Sun, Lionel Delphin-Poulat, Christèle Tarnec and Anastasia Shimorina . . . . .                                                                   | 7997 |
| <i>ConCISE: Confidence-guided Compression in Step-by-step Efficient Reasoning</i>                                                                   |      |
| Ziqing Qiao, Yongheng Deng, Jiali Zeng, Dong Wang, Lai Wei, Guanbo Wang, Fandong Meng, Jie Zhou, Ju Ren and Yaoyue Zhang . . . . .                  | 8022 |
| <i>Layer-Aware Representation Filtering: Purifying Finetuning Data to Preserve LLM Safety Alignment</i>                                             |      |
| Hao Li, Lijun Li, Zhenghao Lu, Xianyi Wei, Rui Li, Jing Shao and Lei Sha . . . . .                                                                  | 8042 |
| <i>Cross-domain Rumor Detection via Test-Time Adaptation and Large Language Models</i>                                                              |      |
| Yuxia Gong, Shuguo Hu and Huaiwen Zhang . . . . .                                                                                                   | 8063 |
| <i>MLWQ: Efficient Small Language Model Deployment via Multi-Level Weight Quantization</i>                                                          |      |
| Chun Hu, Junhui He, Shangyu Wu, Yuxin He, Chun Jason Xue and Qingan Li . . . . .                                                                    | 8079 |

|                                                                                                                                     |      |
|-------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>ToDi: Token-wise Distillation via Fine-Grained Divergence Control</i>                                                            |      |
| Seongryong Jung, Suwan Yoon, DongGeon Kim and Hwanhee Lee .....                                                                     | 8090 |
| <i>RethinkMCTS: Refining Erroneous Thoughts in Monte Carlo Tree Search for Code Generation</i>                                      |      |
| Qingyao Li, Wei Xia, Xinyi Dai, Kounianhua Du, Weiwen Liu, Yasheng Wang, Ruiming Tang, Yong Yu and Weinan Zhang .....               | 8104 |
| <i>Probing for Arithmetic Errors in Language Models</i>                                                                             |      |
| Yucheng Sun, Alessandro Stolfo and Mrinmaya Sachan .....                                                                            | 8123 |
| <i>NILE: Internal Consistency Alignment in Large Language Models</i>                                                                |      |
| Minda Hu, Qiyuan Zhang, Yufei Wang, Bowei He, Hongru Wang, Jingyan Zhou, Liangyou Li, Yasheng Wang, Chen Ma and Irwin King .....    | 8141 |
| <i>Mining the Past with Dual Criteria: Integrating Three types of Historical Information for Context-aware Event Forecasting</i>    |      |
| Rong Ma, Lei Wang, Yating Yang, Bo Ma, Rui Dong, Fengyi Yang, Ahtamjan Ahmat, Kaiwen Lu and Xinyue Wang .....                       | 8160 |
| <i>RAGfreee: Building Contextual Reward Models for Retrieval-Augmented Generation</i>                                               |      |
| Andrei Catalin Coman, Ionut Teodor Sorodoc, Leonardo F. R. Ribeiro, Bill Byrne, James Henderson and Adrià de Gispert .....          | 8176 |
| <i>Large Language Models Discriminate Against Speakers of German Dialects</i>                                                       |      |
| Minh Duc Bui, Carolin Holtermann, Valentin Hofmann, Anne Lauscher and Katharina von der Wense .....                                 | 8224 |
| <i>Uncovering Argumentative Flow: A Question-Focus Discourse Structuring Framework</i>                                              |      |
| Yini Wang, Xian Zhou, Shengan Zheng, Linpeng Huang, Zhunchen Luo, Wei Luo and Xiaoying Bai .....                                    | 8253 |
| <i>AbsVis – Benchmarking How Humans and Vision-Language Models See Abstract Concepts in Images</i>                                  |      |
| Tarun Tater, Diego Frassinelli and Sabine Schulte im Walde .....                                                                    | 8272 |
| <i>A Rigorous Evaluation of LLM Data Generation Strategies for Low-Resource Languages</i>                                           |      |
| Tatiana Anikina, Jan Cegin, Jakub Simko and Simon Ostermann .....                                                                   | 8294 |
| <i>Alignment with Fill-In-the-Middle for Enhancing Code Generation</i>                                                              |      |
| Houxing Ren, Zimu Lu, Weikang Shi, Haotian Hou, Yunqiao Yang, Ke Wang, Aojun Zhou, Junting Pan, Mingjie Zhan and Hongsheng Li ..... | 8316 |
| <i>A Middle Path for On-Premises LLM Deployment: Preserving Privacy Without Sacrificing Model Confidentiality</i>                   |      |
| Hanbo Huang, Yihan Li, Bowen Jiang, Bo Jiang, Lin Liu, Zhuotao Liu, Ruoyu Sun and Shiyu Liang .....                                 | 8333 |
| <i>Variance Sensitivity Induces Attention Entropy Collapse and Instability in Transformers</i>                                      |      |
| Jonghyun Hong and Sungyoon Lee .....                                                                                                | 8372 |
| <i>X-FLoRA: Cross-modal Federated Learning with Modality-expert LoRA for Medical VQA</i>                                            |      |
| Min Hyuk Kim, Changheon Kim and Seok Bong Yoo .....                                                                                 | 8391 |
| <i>Robust Native Language Identification through Agentic Decomposition</i>                                                          |      |
| Ahmet Yavuz Uluslu, Tannon Kew, Tilia Ellendorff, Gerold Schneider and Rico Sennrich ..                                             | 8410 |

|                                                                                                                                                                              |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>ConsistentChat: Building Skeleton-Guided Consistent Multi-Turn Dialogues for Large Language Models from Scratch</i>                                                       |      |
| Jiawei Chen, Xinyan Guan, Qianhao Yuan, Mo Guozhao, Weixiang Zhou, Yaojie Lu, Hongyu Lin, Ben He, Le Sun and Xianpei Han . . . . .                                           | 8427 |
| <i>Does Acceleration Cause Hidden Instability in Vision Language Models? Uncovering Instance-Level Divergence Through a Large-Scale Empirical Study</i>                      |      |
| Yizheng Sun, Hao Li, Chang Xu, Hongpeng Zhou, Chenghua Lin, Riza Batista-Navarro and Jingyuan Sun . . . . .                                                                  | 8454 |
| <i>When Annotators Disagree, Topology Explains: Mapper, a Topological Tool for Exploring Text Embedding Geometry and Ambiguity</i>                                           |      |
| Nisrine Rair, Alban Goupil, Valeriu Vrabie and Emmanuel Chochoy . . . . .                                                                                                    | 8469 |
| <i>Self-Critique and Refinement for Faithful Natural Language Explanations</i>                                                                                               |      |
| Yingming Wang and Pepa Atanasova . . . . .                                                                                                                                   | 8493 |
| <i>The Psychology of Falsehood: A Human-Centric Survey of Misinformation Detection</i>                                                                                       |      |
| Arghodeep Nandi, Megha Sundriyal, Euna Mehnaz Khan, Jikai Sun, Emily K. Vraga, Jaideep Srivastava and Tanmoy Chakraborty . . . . .                                           | 8520 |
| <i>SEAL: Structure and Element Aware Learning Improves Long Structured Document Retrieval</i>                                                                                |      |
| Xinhao Huang, Zhibo Ren, Yipeng Yu, Ying Zhou, Zulong Chen and Zeyi Wen . . . . .                                                                                            | 8538 |
| <i>AnchorAttention: Difference-Aware Sparse Attention with Stripe Granularity</i>                                                                                            |      |
| Yu Zhang, Dong Guo, Fang Wu, Guoliang Zhu, Dian Ding and Yiming Zhang . . . . .                                                                                              | 8549 |
| <i>Attacks by Content: Automated Fact-checking is an AI Security Issue</i>                                                                                                   |      |
| Michael Sejr Schlichtkrull . . . . .                                                                                                                                         | 8562 |
| <i>MUZO: Leveraging Multiple Queries and Momentum for Zeroth-Order Fine-Tuning of Large Language Models</i>                                                                  |      |
| Yuezhang Peng, Yuxin Liu, Fei Wen and Xie Chen . . . . .                                                                                                                     | 8578 |
| <i>Your Language Model Can Secretly Write Like Humans: Contrastive Paraphrase Attacks on LLM-Generated Text Detectors</i>                                                    |      |
| Hao Fang, Jiawei Kong, Tianqu Zhuang, Yixiang Qiu, Kuofeng Gao, Bin Chen, Shu-Tao Xia, Yaowei Wang and Min Zhang . . . . .                                                   | 8597 |
| <i>Will It Still Be True Tomorrow? Multilingual Evergreen Question Classification to Improve Trustworthy QA</i>                                                              |      |
| Sergey Pletenev, Maria Marina, Nikolay Ivanov, Daria Galimzianova, Nikita Krayko, Mikhail Salnikov, Vasily Konovalov, Alexander Panchenko and Viktor Moskvoretskii . . . . . | 8615 |
| <i>Steering Language Models in Multi-Token Generation: A Case Study on Tense and Aspect</i>                                                                                  |      |
| Alina Klerings, Jannik Brinkmann, Daniel Ruffinelli and Simone Paolo Ponzetto . . . . .                                                                                      | 8633 |
| <i>DocReRank: Single-Page Hard Negative Query Generation for Training Multi-Modal RAG Rerankers</i>                                                                          |      |
| Navve Wasserman, Oliver Heinimann, Yuval Golbari, Tal Zimbalist, Eli Schwartz and Michal Irani . . . . .                                                                     | 8652 |
| <i>Reason to Rote: Rethinking Memorization in Reasoning</i>                                                                                                                  |      |
| Yupei Du, Philipp Mondorf, Silvia Casola, Yuekun Yao, Robert Litschko and Barbara Plank . . . . .                                                                            | 8671 |
| <i>VELA: An LLM-Hybrid-as-a-Judge Approach for Evaluating Long Image Captions</i>                                                                                            |      |
| Kazuki Matsuda, Yuiga Wada, Shinnosuke Hirano, Seitaro Otsuki and Komei Sugiura . . . . .                                                                                    | 8692 |



|                                                                                                                                                                                                                                                                                                              |      |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>LLM-Independent Adaptive RAG: Let the Question Speak for Itself</i>                                                                                                                                                                                                                                       |      |
| Maria Marina, Nikolay Ivanov, Sergey Pletenev, Mikhail Salnikov, Daria Galimzianova, Nikita Krayko, Vasily Kononov, Alexander Panchenko and Viktor Moskvoretskii                                                                                                                                             | 8709 |
| <i>TurnBack: A Geospatial Route Cognition Benchmark for Large Language Models through Reverse Route</i>                                                                                                                                                                                                      |      |
| Hongyi Luo, Qing Cheng, Daniel Matos, Hari Krishna Gadi, Yanfeng Zhang, Lu Liu, Yongliang Wang, Niclas Zeller, Daniel Cremers and Liqiu Meng                                                                                                                                                                 | 8722 |
| <i>Certainty in Uncertainty: Reasoning over Uncertain Knowledge Graphs with Statistical Guarantees</i>                                                                                                                                                                                                       |      |
| Yuqicheng Zhu, Jingcheng Wu, Yizhen Wang, Hongkuan Zhou, Jiaoyan Chen, Evgeny Kharlamov and Steffen Staab                                                                                                                                                                                                    | 8742 |
| <i>Beyond Seen Data: Improving KBQA Generalization Through Schema-Guided Logical Form Generation</i>                                                                                                                                                                                                         |      |
| Shengxiang Gao, Jey Han Lau and Jianzhong Qi                                                                                                                                                                                                                                                                 | 8765 |
| <i>A Training-Free Length Extrapolation Approach for LLMs: Greedy Attention Logit Interpolation</i>                                                                                                                                                                                                          |      |
| Yan Li, Tianyi Zhang, Zechuan Li and Caren Han                                                                                                                                                                                                                                                               | 8785 |
| <i>Taming Text-to-Image Synthesis for Novices: User-centric Prompt Generation via Multi-turn Guidance</i>                                                                                                                                                                                                    |      |
| Yilun Liu, Mingui HE, Feiyu Yao, Yuhe Ji, Shimin Tao, Jingzhou DU, Justin Li, Jian Gao, Zhang Li, Hao Yang, Boxing Chen and Osamu Yoshie                                                                                                                                                                     | 8806 |
| <i>We Need to Measure Data Diversity in NLP — Better and Broader</i>                                                                                                                                                                                                                                         |      |
| Dong Nguyen and Esther Ploeger                                                                                                                                                                                                                                                                               | 8824 |
| <i>Sheaf Discovery with Joint Computation Graph Pruning and Flexible Granularity</i>                                                                                                                                                                                                                         |      |
| Lei Yu, Jingcheng Niu, Zining Zhu, Xi Chen and Gerald Penn                                                                                                                                                                                                                                                   | 8834 |
| <i>Hierarchical Bracketing Encodings Work for Dependency Graphs</i>                                                                                                                                                                                                                                          |      |
| Ana Ezquerro, Carlos Gómez-Rodríguez and David Vilares                                                                                                                                                                                                                                                       | 8850 |
| <i>Multimodal Fine-grained Context Interaction Graph Modeling for Conversational Speech Synthesis</i>                                                                                                                                                                                                        |      |
| Zhenqi Jia, Rui Liu, Berrak Sisman and Haizhou Li                                                                                                                                                                                                                                                            | 8864 |
| <i>Judging Quality Across Languages: A Multilingual Approach to Pretraining Data Filtering with Language Models</i>                                                                                                                                                                                          |      |
| Mehdi Ali, Manuel Brack, Max Lübbering, Elias Wendt, Abbas Goher Khan, Richard Rutmann, Alex Jude, Maurice Kraus, Alexander Arno Weber, Felix Stollenwerk, David Kaczér, Florian Mai, Lucie Flek, Rafet Sifa, Nicolas Flores-Herr, Joachim Koehler, Patrick Schramowski, Michael Fromm and Kristian Kersting | 8871 |
| <i>Conditional [MASK] Discrete Diffusion Language Model</i>                                                                                                                                                                                                                                                  |      |
| Hyukhun Koh, Minha Jhang, Dohyung Kim, Sangmook Lee and Kyomin Jung                                                                                                                                                                                                                                          | 8911 |
| <i>Language-Guided Temporal Token Pruning for Efficient VideoLLM Processing</i>                                                                                                                                                                                                                              |      |
| Yogesh Kumar                                                                                                                                                                                                                                                                                                 | 8936 |
| <i>A Fully Probabilistic Perspective on Large Language Model Unlearning: Evaluation and Optimization</i>                                                                                                                                                                                                     |      |
| Anda Cheng, Wei Huang and Yinggui Wang                                                                                                                                                                                                                                                                       | 8944 |
| <i>IJET: Efficient Numerical Transformer via Implicit Iterative Euler Method</i>                                                                                                                                                                                                                             |      |
| Xinyu Liu, Bei Li, Jiahao Liu, Junhao Ruan, Kechen Jiao, Hongyin Tang, Jingang Wang, Tong Xiao and JingBo Zhu                                                                                                                                                                                                | 8956 |

|                                                                                                                                                  |      |
|--------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>WebEvolver: Enhancing Web Agent Self-Improvement with Co-evolving World Model</i>                                                             |      |
| Tianqing Fang, Hongming Zhang, Zhisong Zhang, Kaixin Ma, Wenhao Yu, Haitao Mi and Dong Yu.....                                                   | 8971 |
| <i>Leveraging Semantic Triples for Private Document Generation with Local Differential Privacy Guarantees</i>                                    |      |
| Stephen Meisenbacher, Maulik Chevli and Florian Matthes.....                                                                                     | 8988 |
| <i>HVGuard: Utilizing Multimodal Large Language Models for Hateful Video Detection</i>                                                           |      |
| Yiheng Jing, Mingming Zhang, Yong Zhuang, Jiacheng Guo, Juan Wang, Xiaoyang Xu, Wenzhe Yi, Keyan Guo and Hongxin Hu.....                         | 9005 |
| <i>Accelerate Parallelizable Reasoning via Parallel Decoding within One Sequence</i>                                                             |      |
| Yijiong Yu, Wei Wang, Ran Chen and Ji Pei.....                                                                                                   | 9019 |
| <i>SlideCoder: Layout-aware RAG-enhanced Hierarchical Slide Generation from Design</i>                                                           |      |
| Wenxin Tang, Jingyu Xiao, Wenxuan Jiang, Xi Xiao, Yuhang Wang, Xuxin Tang, Qing Li, Yuehe Ma, Junliang Liu, Shisong Tang and Michael R. Lyu..... | 9027 |
| <i>LLM-OREF: An Open Relation Extraction Framework Based on Large Language Models</i>                                                            |      |
| Hongyao Tu, Liang Zhang, Yujie Lin, Xin Lin, Haibo Zhang, Long Zhang and Jinsong Su..                                                            | 9052 |
| <i>Ambiguity Awareness Optimization: Towards Semantic Disambiguation for Direct Preference Optimization</i>                                      |      |
| Jian Li, Shenglin Yin, Yujia Zhang, Alan Zhao, Xi Chen, Xiaohui Zhou and Pengfei Xu ...                                                          | 9065 |
| <i>Improving Multilingual Retrieval-Augmented Language Models through Dialectic Reasoning Arguments</i>                                          |      |
| Leonardo Ranaldi, Federico Ranaldi, Fabio Massimo Zanzotto, Barry Haddow and Alexandra Birch.....                                                | 9076 |
| <i>Predicate-Guided Generation for Mathematical Reasoning</i>                                                                                    |      |
| Jiajun Chen and Yik-Cheung Tam.....                                                                                                              | 9098 |
| <i>ComplexTempQA: A 100m Dataset for Complex Temporal Question Answering</i>                                                                     |      |
| Raphael Gruber, Abdelrahman Abdallah, Michael Färber and Adam Jatowt.....                                                                        | 9112 |
| <i>ViDoRAG: Visual Document Retrieval-Augmented Generation via Dynamic Iterative Reasoning Agents</i>                                            |      |
| Qiuchen Wang, Ruixue Ding, Zehui Chen, Weiqi Wu, Shihang Wang, Pengjun Xie and Feng Zhao                                                         | 9125 |
| <i>IndoSafety: Culturally Grounded Safety for LLMs in Indonesian Languages</i>                                                                   |      |
| Muhammad Falensi Azmi, Muhammad Dehan Al Kautsar, Alfian Farizki Wicaksono and Fajri Koto.....                                                   | 9147 |
| <i>Can LLMs Help You at Work? A Sandbox for Evaluating LLM Agents in Enterprise Environments</i>                                                 |      |
| Harsh Vishwakarma, Ankush Agarwal, Ojas Patil, Chaitanya Devaguptapu and Mahesh Chandran                                                         | 9179 |
| <i>Steering LLM Reasoning Through Bias-Only Adaptation</i>                                                                                       |      |
| Viacheslav Sinii, Alexey Gorbatovski, Artem Cherepanov, Boris Shaposhnikov, Nikita Balagansky and Daniil Gavrilov.....                           | 9214 |
| <i>VLASCD: A Visual Language Action Model for Simultaneous Chatting and Decision Making</i>                                                      |      |
| Zuojin Tang, Bin Hu, Chenyang Zhao, De Ma, Gang Pan and Bin Liu.....                                                                             | 9224 |



|                                                                                                                                                                         |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>M-LongDoc: A Benchmark For Multimodal Super-Long Document Understanding And A Retrieval-Aware Tuning Framework</i>                                                   |      |
| Yew Ken Chia, Liying Cheng, Hou Pong Chan, Maojia Song, Chaoqun Liu, Mahani Aljunied, Soujanya Poria and Lidong Bing .....                                              | 9245 |
| <i>Look Again, Think Slowly: Enhancing Visual Reflection in Vision-Language Models</i>                                                                                  |      |
| Pu Jian, Junhong Wu, Wei Sun, Chen Wang, Shuo Ren and Jiajun Zhang .....                                                                                                | 9263 |
| <i>FB-Bench: A Fine-Grained Multi-Task Benchmark for Evaluating LLMs' Responsiveness to Human Feedback</i>                                                              |      |
| Youquan Li, Miao Zheng, Fan Yang, Guosheng Dong, Bin Cui, Weipeng Chen, Zenan Zhou and Wentao Zhang .....                                                               | 9283 |
| <i>HYDRA: A Multi-Head Encoder-only Architecture for Hierarchical Text Classification</i>                                                                               |      |
| Fabian Karl and Ansgar Scherp .....                                                                                                                                     | 9304 |
| <i>CARD: Cross-modal Agent Framework for Generative and Editable Residential Design</i>                                                                                 |      |
| Pengyu Zeng, Jun Yin, Miao Zhang, Yuqin Dai, Jizhizi Li, ZhanXiang Jin and Shuai Lu ...                                                                                 | 9316 |
| <i>DrDiff: Dynamic Routing Diffusion with Hierarchical Attention for Breaking the Efficiency-Quality Trade-off</i>                                                      |      |
| Jusheng Zhang, Yijia Fan, Kaitong Cai, Zimeng Huang, Xiaofei Sun, Jian Wang, Chengpei Tang and Keze Wang .....                                                          | 9332 |
| <i>FaST: Feature-aware Sampling and Tuning for Personalized Preference Alignment with Limited Data</i>                                                                  |      |
| Thibaut Thonet, Germán Kruszewski, Jos Rozen, Pierre Erbacher and Marc Dymetman ...                                                                                     | 9353 |
| <i>On LLM-Based Scientific Inductive Reasoning Beyond Equations</i>                                                                                                     |      |
| Brian S. Lin, Jiaxin Yuan, Zihan Zhou, Shouli Wang, Shuo Wang, Cunliang Kong, Qi Shi, Yuxuan Li, Liner Yang, Zhiyuan Liu and Maosong Sun .....                          | 9383 |
| <i>SPECS: Specificity-Enhanced CLIP-Score for Long Image Caption Evaluation</i>                                                                                         |      |
| Xiaofu Chen, Israfel Salazar and Yova Kementchedjheva .....                                                                                                             | 9407 |
| <i>LM-Searcher: Cross-domain Neural Architecture Search with LLMs via Unified Numerical Encoding</i>                                                                    |      |
| Yuxuan Hu, Jihao Liu, Ke Wang, Jinliang Zheng, Weikang Shi, Manyuan Zhang, Qi Dou, Rui Liu, Aojun Zhou and Hongsheng Li .....                                           | 9420 |
| <i>Does quantization affect models' performance on long-context tasks?</i>                                                                                              |      |
| Anmol Mekala, Anirudh Atmakuru, Yixiao Song, Marzena Karpinska and Mohit Iyyer ...                                                                                      | 9434 |
| <i>Token-Aware Editing of Internal Activations for Large Language Model Alignment</i>                                                                                   |      |
| Tianbo Wang, Yuqing Ma, Kewei Liao, Chengzhao Yang, Zhange Zhang, Jiakai Wang and Xian-long Liu .....                                                                   | 9483 |
| <i>Bitune: Leveraging Bidirectional Attention to Improve Decoder-Only LLMs</i>                                                                                          |      |
| Dawid Jan Kopiczko, Tijmen Blankevoort and Yuki M Asano .....                                                                                                           | 9522 |
| <i>Disambiguation in Conversational Question Answering in the Era of LLMs and Agents: A Survey</i>                                                                      |      |
| Mehrab Tanjim, Yeonjun In, Xiang Chen, Victor Burszty, Ryan A. Rossi, Sungchul Kim, Guang-Jie Ren, Vaishnavi Muppala, Shun Jiang, Yongsung Kim and Chanyoung Park ..... | 9549 |
| <i>Plan Dynamically, Express Rhetorically: A Debate-Driven Rhetorical Framework for Argumentative Writing</i>                                                           |      |
| Xueguan Zhao, Wenpeng Lu, Chaoqun Zheng, Weiyu Zhang, Jiasheng Si and Deyu Zhou ..                                                                                      | 9563 |

|                                                                                                                                                                                                                                                                                                                                                                   |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <i>TCPO: Thought-Centric Preference Optimization for Effective Embodied Decision-making</i>                                                                                                                                                                                                                                                                       |      |
| Kechen Jiao, Zhirui Fang, Jiahao Liu, Bei Li, Qifan Wang, Xinyu Liu, Junhao Ruan, Zhongjian Qiao, Yifan Zhu, Yaxin Xu, Jingang Wang and Xiu Li .....                                                                                                                                                                                                              | 9586 |
| <i>Reimagining Safety Alignment with An Image</i>                                                                                                                                                                                                                                                                                                                 |      |
| Yifan Xia, Guorui Chen, Wenqian Yu, Zhijiang Li, Philip Torr and Jindong Gu .....                                                                                                                                                                                                                                                                                 | 9601 |
| <i>Generative or Discriminative? Revisiting Text Classification in the Era of Transformers</i>                                                                                                                                                                                                                                                                    |      |
| Siva Rajesh Kasa, Karan Gupta, Sumegh Roychowdhury, Ashutosh Kumar, Yaswanth Biruduraju, Santhosh Kumar Kasa, Pattisapu Nikhil Priyatam, Arindam Bhattacharya, Shailendra Agarwal and Vijay Huddar .....                                                                                                                                                          | 9616 |
| <i>Visual Contextual Attack: Jailbreaking MLLMs with Image-Driven Context Injection</i>                                                                                                                                                                                                                                                                           |      |
| Miao Ziqi, Yi Ding, Lijun Li and Jing Shao .....                                                                                                                                                                                                                                                                                                                  | 9639 |
| <i>Can Large Language Models Win the International Mathematical Games?</i>                                                                                                                                                                                                                                                                                        |      |
| Alessio Cocchieri, Luca Ragazzi, Giuseppe Tagliavini, Lorenzo Tordi, Antonella Carbonaro and Gianluca Moro .....                                                                                                                                                                                                                                                  | 9657 |
| <i>CodeArena: Evaluating and Aligning CodeLLMs on Human Preference</i>                                                                                                                                                                                                                                                                                            |      |
| Jian Yang, Jiaxi Yang, Wei Zhang, Jin Ke, Yibo Miao, Lei Zhang, Liquan Yang, Zeyu Cui, Yichang Zhang, Zhoujun Li, Binyuan Hui and Junyang Lin .....                                                                                                                                                                                                               | 9684 |
| <i>Language models can learn implicit multi-hop reasoning, but only if they have lots of training data</i>                                                                                                                                                                                                                                                        |      |
| Yuekun Yao, Yupei Du, Dawei Zhu, Michael Hahn and Alexander Koller .....                                                                                                                                                                                                                                                                                          | 9696 |
| <i>UniversalCEFR: Enabling Open Multilingual Research on Language Proficiency Assessment</i>                                                                                                                                                                                                                                                                      |      |
| Joseph Marvin Imperial, Abdullah Barayan, Regina Stodden, Rodrigo Wilkens, Ricardo Muñoz Sánchez, Lingyun Gao, Melissa Torgbi, Dawn Knight, Gail Forey, Reka R. Jablonkai, Ekaterina Kochmar, Robert Joshua Reynolds, Eugénio Ribeiro, Horacio Saggion, Elena Volodina, Sowmya Vajjala, Thomas François, Fernando Alva-Manchego and Harish Tayyar Madabushi ..... | 9715 |
| <i>CROP: Contextual Region-Oriented Visual Token Pruning</i>                                                                                                                                                                                                                                                                                                      |      |
| Jiawei Guo, Feifei Zhai, Pu Jian, Qianrun Wei and Yu Zhou .....                                                                                                                                                                                                                                                                                                   | 9768 |
| <i>CR4-NarrEmote: An Open Vocabulary Dataset of Narrative Emotions Derived Using Citizen Science</i>                                                                                                                                                                                                                                                              |      |
| Andrew Piper and Robert Budac .....                                                                                                                                                                                                                                                                                                                               | 9785 |
| <i>XQuant: Achieving Ultra-Low Bit KV Cache Quantization with Cross-Layer Compression</i>                                                                                                                                                                                                                                                                         |      |
| Haoqi Yang, Yao Yao, Zuchao Li, Baoyuan Qi, Liu Guoming and Hai Zhao .....                                                                                                                                                                                                                                                                                        | 9797 |
| <i>DINT Transformer</i>                                                                                                                                                                                                                                                                                                                                           |      |
| Yueyang Cang, Yuhang Liu, Xiaoteng Zhang, Erlu Zhao and Li Shi .....                                                                                                                                                                                                                                                                                              | 9813 |
| <i>ICR: Iterative Clarification and Rewriting for Conversational Search</i>                                                                                                                                                                                                                                                                                       |      |
| Zhiyu Cao, Peifeng Li and Qiaoming Zhu .....                                                                                                                                                                                                                                                                                                                      | 9822 |
| <i>Pre-training CLIP against Data Poisoning with Optimal Transport-based Matching and Alignment</i>                                                                                                                                                                                                                                                               |      |
| Tong Zhang, Kuofeng Gao, Jiawang Bai, Leo Yu Zhang, Xin Yin, Zonghui Wang, Shouling Ji and Wenzhi Chen .....                                                                                                                                                                                                                                                      | 9837 |
| <i>Similarity = Value? Consultation Value-Assessment and Alignment for Personalized Search</i>                                                                                                                                                                                                                                                                    |      |
| Weicong Qin, Yi Xu, Weijie Yu, Teng Shi, Chenglei Shen, Ming He, Jianping Fan, Xiao Zhang and Jun Xu .....                                                                                                                                                                                                                                                        | 9851 |
| <i>RTQA : Recursive Thinking for Complex Temporal Knowledge Graph Question Answering with Large Language Models</i>                                                                                                                                                                                                                                               |      |
| Zhaoyan Gong, Juan Li, Zhiqiang Liu, Lei Liang, Huajun Chen and Wen Zhang .....                                                                                                                                                                                                                                                                                   | 9865 |

|                                                                                                                                           |       |
|-------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Not All Parameters Are Created Equal: Smart Isolation Boosts Fine-Tuning Performance</i>                                               |       |
| Yao Wang, Di Liang and Minlong Peng .....                                                                                                 | 9883  |
| <i>AI Knows Where You Are: Exposure, Bias, and Inference in Multimodal Geolocation with KoreaGEO</i>                                      |       |
| Xiaonan Wang, Bo Shao and Hansaem Kim .....                                                                                               | 9898  |
| <i>CAT: Causal Attention Tuning For Injecting Fine-grained Causal Knowledge into Large Language Models</i>                                |       |
| Kairong Han, Wenshuo Zhao, Ziyu Zhao, Ye Jun Jian, Lujia Pan and Kun Kuang .....                                                          | 9916  |
| <i>Enhancing LLM Text Detection with Retrieved Contexts and Logits Distribution Consistency</i>                                           |       |
| Zhaoheng Huang, Yutao Zhu, Ji-Rong Wen and Zhicheng Dou .....                                                                             | 9934  |
| <i>Measuring Chain of Thought Faithfulness by Unlearning Reasoning Steps</i>                                                              |       |
| Martin Tutek, Fateme Hashemi Chaleshtori, Ana Marasovic and Yonatan Belinkov .....                                                        | 9947  |
| <i>Stop Looking for “Important Tokens” in Multimodal Language Models: Duplication Matters More</i>                                        |       |
| Zichen Wen, Yifeng Gao, Shaobo Wang, Junyuan Zhang, Qintong Zhang, Weijia Li, Conghui He and Linfeng Zhang .....                          | 9973  |
| <i>AgentPro: Enhancing LLM Agents with Automated Process Supervision</i>                                                                  |       |
| Yuchen Deng, Shichen Fan, Naibo Wang, Xinkui Zhao and See-Kiong Ng .....                                                                  | 9993  |
| <i>PORTS: Preference-Optimized Retrievers for Tool Selection with Large Language Models</i>                                               |       |
| Lorenzo Molfetta, Giacomo Frisoni, Nicolò Monaldini and Gianluca Moro .....                                                               | 10019 |
| <i>MusKGC: A Flexible Multi-source Knowledge Enhancement Framework for Open-World Knowledge Graph Completion</i>                          |       |
| Xin Song, Liu Haiyan, Haiyang Wang, Ye Wang, Kai Chen and Bin Zhou .....                                                                  | 10043 |
| <i>Towards Transferable Personality Representation Learning based on Triplet Comparisons and Its Applications</i>                         |       |
| Kai Tang, Rui Wang, Renyu Zhu, Minmin Lin, Xiao Ding, Tangjie Lv, Changjie Fan, Runze Wu and Haobo Wang .....                             | 10062 |
| <i>Reshaping Representation Space to Balance the Safety and Over-rejection in Large Audio Language Models</i>                             |       |
| Hao Yang, Lizhen Qu, Ehsan Shareghi and Gholamreza Haffari .....                                                                          | 10079 |
| <i>Benchmarking Large Language Models Under Data Contamination: A Survey from Static to Dynamic Evaluation</i>                            |       |
| Simin Chen, Yiming Chen, Zexin Li, Yifan Jiang, Zhongwei Wan, Yixin He, Dezhi Ran, Tianle Gu, Haizhou Li, Tao Xie and Baishakhi Ray ..... | 10092 |
| <i>FinTrust: A Comprehensive Benchmark of Trustworthiness Evaluation in Finance Domain</i>                                                |       |
| Tiansheng Hu, Tongyan Hu, Liuyang Bai, Yilun Zhao, Arman Cohan and Chen Zhao .....                                                        | 10111 |
| <i>RecGPT: A Foundation Model for Sequential Recommendation</i>                                                                           |       |
| Yangqin Jiang, Xubin Ren, Lianghao Xia, Da Luo, Kangyi Lin and Chao Huang .....                                                           | 10141 |
| <i>Towards Holistic Evaluation of Large Audio-Language Models: A Comprehensive Survey</i>                                                 |       |
| Chih-Kai Yang, Neo S. Ho and Hung-yi Lee .....                                                                                            | 10156 |
| <i>Train One Sparse Autoencoder Across Multiple Sparsity Budgets to Preserve Interpretability and Accuracy</i>                            |       |
| Nikita Balagansky, Yaroslav Aksenov, Daniil Laptev, Vadim Kurochkin, Gleb Gerasimov, Nikita Koriagin and Daniil Gavrilov .....            | 10183 |

|                                                                                                                                                                   |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Learn and Unlearn: Addressing Misinformation in Multilingual LLMs</i>                                                                                          |       |
| TaiMing Lu and Philipp Koehn .....                                                                                                                                | 10192 |
| <i>PRISM: Efficient Long-Range Reasoning With Short-Context LLMs</i>                                                                                              |       |
| Dulhan Jayalath, James Bradley Wendt, Nicholas Monath, Sandeep Tata and Beliz Gunel .                                                                             | 10208 |
| <i>Augmenting Multi-Agent Communication with State Delta Trajectory</i>                                                                                           |       |
| Yichen Tang, Weihang Su, Yujia Zhou, Yiqun Liu, Min Zhang, Shaoping Ma and Qingyao Ai                                                                             | 10231 |
| <i>SAEs Are Good for Steering – If You Select the Right Features</i>                                                                                              |       |
| Dana Arad, Aaron Mueller and Yonatan Belinkov .....                                                                                                               | 10253 |
| <i>CoBA: Counterbias Text Augmentation for Mitigating Various Spurious Correlations via Semantic Triples</i>                                                      |       |
| Kyohoon Jin, Juhwan Choi, JungMin Yun, Junho Lee, Soojin Jang and YoungBin Kim . . .                                                                              | 10272 |
| <i>Layered Insights: Generalizable Analysis of Human Authorial Style by Leveraging All Transformer Layers</i>                                                     |       |
| Milad Alshomary, Nikhil Reddy Varimalla, Vishal Anand, Smaranda Muresan and Kathleen McKeown .....                                                                | 10291 |
| <i>When Long Helps Short: How Context Length in Supervised Fine-tuning Affects Behavior of Large Language Models</i>                                              |       |
| Yingming Zheng, Hanqi Li, Kai Yu and Lu Chen .....                                                                                                                | 10305 |
| <i>A Case Against Implicit Standards: Homophone Normalization in Machine Translation for Languages that use the Ge'ez Script.</i>                                 |       |
| Hellina Hailu Nigatu, Atnafu Lambebo Tonja, Henok Biadgign Ademtew, Hizkiel Mitiku Alemayehu, Negasi Haile Abadi, Tadesse Destaw Belay and Seid Muhie Yimam ..... | 10321 |
| <i>Evaluating Language Translation Models by Playing Telephone</i>                                                                                                |       |
| Syeda Jannatus Saba and Steven Skiena .....                                                                                                                       | 10333 |
| <i>Doubling Your Data in Minutes: Ultra-fast Tabular Data Generation via LLM-Induced Dependency Graphs</i>                                                        |       |
| Shuo Yang, Zheyu Zhang, Bardh Prenkaj and Gjergji Kasneci .....                                                                                                   | 10349 |
| <i>SPaRC: A Spatial Pathfinding Reasoning Challenge</i>                                                                                                           |       |
| Lars Benedikt Kaesberg, Jan Philip Wahle, Terry Ruas and Bela Gipp .....                                                                                          | 10371 |
| <i>Primus: A Pioneering Collection of Open-Source Datasets for Cybersecurity LLM Training</i>                                                                     |       |
| Yao-Ching Yu, Tsun-Han Chiang, Cheng-Wei Tsai, Chien-Ming Huang and Wen-Kwang Tsao                                                                                | 10403 |
| <i>Bit-Flip Error Resilience in LLMs: A Comprehensive Analysis and Defense Framework</i>                                                                          |       |
| Yuhang Chen, Zhen Tan, Ajay Kumar Jaiswal, Huaizhi Qu, Xinyu Zhao, Qi Lin, Yu Cheng, Andrew Kwong, Zhichao Cao and Tianlong Chen .....                            | 10426 |
| <i>Towards Faithful Natural Language Explanations: A Study Using Activation Patching in Large Language Models</i>                                                 |       |
| Wei Jie Yeo, Ranjan Satapathy and Erik Cambria .....                                                                                                              | 10437 |
| <i>Calibrating LLM Confidence by Probing Perturbed Representation Stability</i>                                                                                   |       |
| Reza Khanmohammadi, Erfan Miah, Mehrsa Mardikoraem, Simerjot Kaur, Ivan Brugere, Charese Smiley, Kundan S Thind and Mohammad M. Ghassemi .....                    | 10460 |

|                                                                                                                                                                     |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>SATER: A Self-Aware and Token-Efficient Approach to Routing and Cascading</i>                                                                                    |       |
| Yuanzhe Shen, Yide Liu, Zisu Huang, Ruicheng Yin, Xiaoqing Zheng and Xuanjing Huang                                                                                 | 10527 |
| <i>DSG-MCTS: A Dynamic Strategy-Guided Monte Carlo Tree Search for Diversified Reasoning in Large Language Models</i>                                               |       |
| Rui Ha, Chaozhuo Li, Rui Pu, Litian Zhang, Xi Zhang and Sen Su                                                                                                      | 10542 |
| <i>CIFLEX: Contextual Instruction Flow for Sub-task Execution in Multi-Turn Interactions with a Single On-Device LLM</i>                                            |       |
| Juntae Lee, Jihwan Bang, Seunghan Yang and Simyung Chang                                                                                                            | 10557 |
| <i>On the Role of Model Prior in Real-World Inductive Reasoning</i>                                                                                                 |       |
| Zhuo Liu, Ding Yu and Hangfeng He                                                                                                                                   | 10572 |
| <i>Viability of Machine Translation for Healthcare in Low-Resourced Languages</i>                                                                                   |       |
| Hellina Hailu Nigatu, Nikita Mehandru, Negasi Haile Abadi, Blen Gebremeskel, Ahmed Alaa and Monojit Choudhury                                                       | 10596 |
| <i>Latent Inter-User Difference Modeling for LLM Personalization</i>                                                                                                |       |
| Yilun Qiu, Tianhao Shi, Xiaoyan Zhao, Fengbin Zhu, Yang Zhang and Fuli Feng                                                                                         | 10611 |
| <i>IG-Pruning: Input-Guided Block Pruning for Large Language Models</i>                                                                                             |       |
| Kangyu Qiao, Shaolei Zhang and Yang Feng                                                                                                                            | 10630 |
| <i>Are Checklists Really Useful for Automatic Evaluation of Generative Tasks?</i>                                                                                   |       |
| Momoka Furuhashi, Kouta Nakayama, Takashi Kodama and Saku Sugawara                                                                                                  | 10642 |
| <i>Measuring the Effect of Disfluency in Multilingual Knowledge Probing Benchmarks</i>                                                                              |       |
| Kirill Semenov and Rico Sennrich                                                                                                                                    | 10666 |
| <i>Knowledge Editing through Chain-of-Thought</i>                                                                                                                   |       |
| Changyue Wang, Weihang Su, Qingyao Ai, Yichen Tang and Yiqun Liu                                                                                                    | 10685 |
| <i>SelfRACG: Enabling LLMs to Self-Express and Retrieve for Code Generation</i>                                                                                     |       |
| Qian Dong, Jia Chen, Qingyao Ai, Hongning Wang, Haitao Li, Yiwu, Yao Hu, Yiqun Liu and Shaoping Ma                                                                  | 10706 |
| <i>Probing Logical Reasoning of MLLMs in Scientific Diagrams</i>                                                                                                    |       |
| Yufei Wang and Adriana Kovashka                                                                                                                                     | 10718 |
| <i>AdamS: Momentum Itself Can Be A Normalizer for LLM Pretraining and Post-training</i>                                                                             |       |
| Huishuai Zhang, Bohan Wang and Luoxin Chen                                                                                                                          | 10731 |
| <i>Demystifying Synthetic Data in LLM Pre-training: A Systematic Study of Scaling Laws, Benefits, and Pitfalls</i>                                                  |       |
| Feiyang Kang, Newsha Ardalani, Michael Kuchnik, Youssef Emad, Mostafa Elhoushi, Shubhabrata Sengupta, Shang-Wen Li, Ramya Raghavendra, Ruoxi Jia and Carole-Jean Wu | 10751 |
| <i>Static or Dynamic: Towards Query-Adaptive Token Selection for Video Question Answering</i>                                                                       |       |
| Yumeng Shi, Quanyu Long and Wenya Wang                                                                                                                              | 10771 |
| <i>DischargeSim: A Simulation Benchmark for Educational Doctor–Patient Communication at Discharge</i>                                                               |       |
| Zonghai Yao, Michael Sun, Won Seok Jang, Sunjae Kwon, Soie Kwon and Hong yu                                                                                         | 10784 |
| <i>Can Vision-Language Models Solve Visual Math Equations?</i>                                                                                                      |       |
| Monjoy Narayan Choudhury, Junling Wang, Yifan Hou and Mrinmaya Sachan                                                                                               | 10811 |

|                                                                                                                                                   |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>From Scores to Steps: Diagnosing and Improving LLM Performance in Evidence-Based Medical Calculations</i>                                      |       |
| Benlu Wang, Iris Xia, Yifan Zhang, Junda Wang, Feiyun Ouyang , Shuo Han, Arman Cohan, Hong yu and Zonghai Yao .....                               | 10821 |
| <i>Bridging External and Parametric Knowledge: Mitigating Hallucination of LLMs with Shared-Private Semantic Synergy in Dual-Stream Knowledge</i> |       |
| Yi Sui, Chaozhuo Li, Chen Zhang, Dawei Song and Qiuchi Li .....                                                                                   | 10846 |
| <i>Deep Associations, High Creativity: A Simple yet Effective Metric for Evaluating Large Language Models</i>                                     |       |
| Ziliang Qiu and Renfen Hu .....                                                                                                                   | 10871 |
| <i>Identifying Unlearned Data in LLMs via Membership Inference Attacks</i>                                                                        |       |
| Advit Deepak, Megan Mou, Jing Huang and Diyi Yang .....                                                                                           | 10885 |
| <i>Feature Extraction and Steering for Enhanced Chain-of-Thought Reasoning in Language Models</i>                                                 |       |
| Zihao Li, Xu Wang, Yuzhe Yang, Ziyu Yao, Haoyi Xiong and Mengnan Du .....                                                                         | 10905 |
| <i>LLMs cannot spot math errors, even when allowed to peek into the solution</i>                                                                  |       |
| Kv Aditya Srivatsa, Kaushal Kumar Maurya and Ekaterina Kochmar .....                                                                              | 10926 |
| <i>Can LLMs be Good Graph Judge for Knowledge Graph Construction?</i>                                                                             |       |
| Haoyu Huang, Chong Chen, Zeang Sheng, Yang Li and Wentao Zhang .....                                                                              | 10941 |
| <i>NeuroAda: Activating Each Neuron's Potential for Parameter-Efficient Fine-Tuning</i>                                                           |       |
| Zhi Zhang, Yixian Shen, Congfeng Cao and Ekaterina Shutova .....                                                                                  | 10961 |
| <i>NileChat: Towards Linguistically Diverse and Culturally Aware LLMs for Local Communities</i>                                                   |       |
| Abdellah El Mekki, Houdaifa Atou, Omer Nacar, Shady Shehata and Muhammad Abdul-Mageed                                                             | 10979 |
| <i>A Computational Simulation of Language Production in First Language Acquisition</i>                                                            |       |
| Yuan Gao and Weiwei Sun .....                                                                                                                     | 11004 |
| <i>Long-Form Information Alignment Evaluation Beyond Atomic Facts</i>                                                                             |       |
| Danna Zheng, Mirella Lapata and Jeff Z. Pan .....                                                                                                 | 11019 |
| <i>Voice of a Continent: Mapping Africa's Speech Technology Frontier</i>                                                                          |       |
| AbdelRahim A. Elmadany, Sang Yun Kwon, Hawau Olamide Toyin, Alcides Alcoba Inciarte, Hanan Aldarmaki and Muhammad Abdul-Mageed .....              | 11040 |
| <i>Cache-Efficient Posterior Sampling for Reinforcement Learning with LLM-Derived Priors Across Discrete and Continuous Domains</i>               |       |
| Ibne Farabi Shihab, Sanjeda Akter and Anuj Sharma .....                                                                                           | 11063 |
| <i>Circuit Complexity Bounds for RoPE-based Transformer Architecture</i>                                                                          |       |
| Bo Chen, Xiaoyu Li, Yingyu Liang, Jiangxuan Long, Zhenmei Shi, Zhao Song and Jiahao Zhang                                                         | 11092 |
| <i>Efficient Unstructured Pruning of Mamba State-Space Models for Resource-Constrained Environments</i>                                           |       |
| Ibne Farabi Shihab, Sanjeda Akter and Anuj Sharma .....                                                                                           | 11110 |
| <i>Towards Infinite-Long Prefix in Transformer</i>                                                                                                |       |
| Yingyu Liang, Zhenmei Shi, Zhao Song and Chiwun Yang .....                                                                                        | 11139 |



|                                                                                                                                                                                        |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>LATTE: Learning to Think with Vision Specialists</i>                                                                                                                                |       |
| Zixian Ma, Jianguo Zhang, Zhiwei Liu, Jieyu Zhang, Juntao Tan, Manli Shu, Juan Carlos Niebles, Shelby Heinecke, Huan Wang, Caiming Xiong, Ranjay Krishna and Silvio Savarese . . . . . | 11204 |
| <i>SUA: Stealthy Multimodal Large Language Model Unlearning Attack</i>                                                                                                                 |       |
| Xianren Zhang, Hui Liu, Delvin Ce Zhang, Xianfeng Tang, Qi He, Dongwon Lee and Suhang Wang . . . . .                                                                                   | 11242 |
| <i>ResFormer: All-Time Reservoir Memory for Long Sequence Classification</i>                                                                                                           |       |
| Hongbo Liu and Jia Xu . . . . .                                                                                                                                                        | 11256 |
| <i>Back Attention: Understanding and Enhancing Multi-Hop Reasoning in Large Language Models</i>                                                                                        |       |
| Zeping Yu, Yonatan Belinkov and Sophia Ananiadou . . . . .                                                                                                                             | 11269 |
| <i>Interdisciplinary Research in Conversation: A Case Study in Computational Morphology for Language Documentation</i>                                                                 |       |
| Enora Rice, Katharina von der Wense and Alexis Palmer . . . . .                                                                                                                        | 11285 |
| <i>Analyzing Uncertainty of LLM-as-a-Judge: Interval Evaluations with Conformal Prediction</i>                                                                                         |       |
| Huanxin Sheng, Xinyi Liu, Hangfeng He, Jieyu Zhao and Jian Kang . . . . .                                                                                                              | 11298 |
| <i>AlphaOne: Reasoning Models Thinking Slow and Fast at Test Time</i>                                                                                                                  |       |
| Junyu Zhang, Runpei Dong, Han Wang, Xuying Ning, Haoran Geng, Peihao Li, Xialin He, Yutong Bai, Jitendra Malik, Saurabh Gupta and Huan Zhang . . . . .                                 | 11341 |
| <i>Dual-Path Dynamic Fusion with Learnable Query for Multimodal Sentiment Analysis</i>                                                                                                 |       |
| Miao Zhou, Lina Yang, Thomas Wu, Dongnan Yang and Xinru Zhang . . . . .                                                                                                                | 11367 |
| <i>CaKE: Circuit-aware Editing Enables Generalizable Knowledge Learners</i>                                                                                                            |       |
| Yunzhi Yao, Jizhan Fang, Jia-Chen Gu, Ningyu Zhang, Shumin Deng, Huajun Chen and Nanyun Peng . . . . .                                                                                 | 11378 |
| <i>DEL-ToM: Inference-Time Scaling for Theory-of-Mind Reasoning via Dynamic Epistemic Logic</i>                                                                                        |       |
| Yuheng Wu, Jianwen Xie, Denghui Zhang and Zhaozhuo Xu . . . . .                                                                                                                        | 11395 |
| <i>Collaborative Beam Search: Enhancing LLM Reasoning via Collective Consensus</i>                                                                                                     |       |
| Yangyifan Xu, Shuo Ren and Jiajun Zhang . . . . .                                                                                                                                      | 11410 |
| <i>Deriving Strategic Market Insights with Large Language Models: A Benchmark for Forward Counterfactual Generation</i>                                                                |       |
| Keane Ong, Rui Mao, Deeksha Varshney, Paul Pu Liang, Erik Cambria and Gianmarco Mengaldo . . . . .                                                                                     | 11423 |
| <i>Towards Statistical Factuality Guarantee for Large Vision-Language Models</i>                                                                                                       |       |
| Zhuohang Li, Chao Yan, Nicholas J Jackson, Wendi Cui, Bo Li, Jiaxin Zhang and Bradley A. Malin . . . . .                                                                               | 11447 |
| <i>Unlearning vs. Obfuscation: Are We Truly Removing Knowledge?</i>                                                                                                                    |       |
| Guangzhi Sun, Potsawee Manakul, Xiao Zhan and Mark Gales . . . . .                                                                                                                     | 11469 |
| <i>Reward-Shifted Speculative Sampling Is An Efficient Test-Time Weak-to-Strong Aligner</i>                                                                                            |       |
| Bolian Li, Yanran Wu, Xinyu Luo and Ruqi Zhang . . . . .                                                                                                                               | 11480 |
| <i>Stimulate the Critical Thinking of LLMs via Debiasing Discussion</i>                                                                                                                |       |
| Ruiyu Xiao, Lei Wu, Yuanxing Liu, Weinan Zhang and Ting Liu . . . . .                                                                                                                  | 11491 |

|                                                                                                                                                      |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Toward Multi-Session Personalized Conversation: A Large-Scale Dataset and Hierarchical Tree Framework for Implicit Reasoning</i>                  |       |
| Xintong Li, Jalend Bantupalli, Ria Dharmani, Yuwei Zhang and Jingbo Shang . . . . .                                                                  | 11505 |
| <i>Improving Instruct Models for Free: A Study on Partial Adaptation</i>                                                                             |       |
| Ozan Irsoy, Pengxiang Cheng, Jennifer L Chen, Daniel Preotiuc-Pietro, Shiyue Zhang and Duccio Pappadopulo . . . . .                                  | 11519 |
| <i>CoMMIT: Coordinated Multimodal Instruction Tuning</i>                                                                                             |       |
| Xintong Li, Junda Wu, Tong Yu, Rui Wang, Yu Wang, Xiang Chen, Jiuxiang Gu, Lina Yao, Julian McAuley and Jingbo Shang . . . . .                       | 11534 |
| <i>Meta-Rewarding Language Models: Self-Improving Alignment with LLM-as-a-Meta-Judge</i>                                                             |       |
| Tianhao Wu, Weizhe Yuan, Olga Golovneva, Jing Xu, Yuandong Tian, Jiantao Jiao, Jason E Weston and Sainbayar Sukhbaatar . . . . .                     | 11549 |
| <i>AnyMAC: Cascading Flexible Multi-Agent Collaboration via Next-Agent Prediction</i>                                                                |       |
| Song Wang, Zhen Tan, Zihan Chen, Shuang Zhou, Tianlong Chen and Jundong Li . . . . .                                                                 | 11567 |
| <i>A Good Plan is Hard to Find: Aligning Models with Preferences is Misaligned with What Helps Users</i>                                             |       |
| Nishant Balepur, Matthew Shu, Yoo Yeon Sung, Seraphina Goldfarb-Tarrant, Shi Feng, Fumeng Yang, Rachel Rudinger and Jordan Lee Boyd-Graber . . . . . | 11580 |
| <i>Words Like Knives: Backstory-Personalized Modeling and Detection of Violent Communication</i>                                                     |       |
| Jocelyn J Shen, Akhila Yerukola, Xuhui Zhou, Cynthia Breazeal, Maarten Sap and Hae Won Park                                                          | 11608 |
| <i>Separate the Wheat from the Chaff: Winnowing Down Divergent Views in Retrieval Augmented Generation</i>                                           |       |
| Song Wang, Zihan Chen, Peng Wang, Zhepei Wei, Zhen Tan, Yu Meng, Cong Shen and Jundong Li . . . . .                                                  | 11627 |
| <i>Cognitive Linguistic Identity Fusion Score (CLIFS): A Scalable Cognition-Informed Approach to Quantifying Identity Fusion from Text</i>           |       |
| Devin R. Wright, Jisun An and Yong-Yeol Ahn . . . . .                                                                                                | 11644 |
| <i>SilVar: Speech-Driven Multimodal Model for Reasoning Visual Question Answering and Object Localization</i>                                        |       |
| Tan-Hanh Pham, Le Hoang Nam, Phu-Vinh Nguyen, Chris Ngo and Truong-Son Hy . . . . .                                                                  | 11675 |
| <i>CEMTM: Contextual Embedding-based Multimodal Topic Modeling</i>                                                                                   |       |
| Amirhossein Abaskohi, Raymond Li, Chuyuan Li, Shafiq Joty and Giuseppe Carenini . . . . .                                                            | 11687 |
| <i>RedHerring Attack: Testing the Reliability of Attack Detection</i>                                                                                |       |
| Jonathan Rusert . . . . .                                                                                                                            | 11705 |
| <i>Modeling Bottom-up Information Quality during Language Processing</i>                                                                             |       |
| Cui Ding, Yanning Yin, Lena Ann Jäger and Ethan Wilcox . . . . .                                                                                     | 11721 |
| <i>Data Drives Unstable Hierarchical Generalization in LMs</i>                                                                                       |       |
| Tian Qin, Naomi Saphra and David Alvarez-Melis . . . . .                                                                                             | 11734 |
| <i>EmoAgent: Assessing and Safeguarding Human-AI Interaction for Mental Health Safety</i>                                                            |       |
| Jiahao Qiu, Yinghui He, Xinzhe Juan, Yimin Wang, Yuhao Liu, Zixin Yao, Yue Wu, Xun Jiang, Ling Yang and Mengdi Wang . . . . .                        | 11753 |



|                                                                                                                                                                                                                   |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Polysemantic Dropout: Conformal OOD Detection for Specialized LLMs</i>                                                                                                                                         |       |
| Ayush Gupta, Ramneet Kaur, Anirban Roy, Adam D. Cobb, Rama Chellappa and Susmit Jha                                                                                                                               | 11769 |
| <i>Facilitating Cognitive Accessibility with LLMs: A Multi-Task Approach to Easy-to-Read Text Generation</i>                                                                                                      |       |
| François Ledoyen, Gaël Dias, Jeremie Pantin, Alexis Lechervy, Fabrice Maurel and Youssef Chahir                                                                                                                   | 11783 |
| <i>D-CoDe: Scaling Image-Pretrained VLMs to Video via Dynamic Compression and Question Decomposition</i>                                                                                                          |       |
| Yiyang Huang, Yizhou Wang and Yun Fu                                                                                                                                                                              | 11810 |
| <i>ReEvalMed: Rethinking Medical Report Evaluation by Aligning Metrics with Real-World Clinical Judgment</i>                                                                                                      |       |
| Ruochen Li, Jun Li, Bailiang Jian, Kun Yuan and Youxiang Zhu                                                                                                                                                      | 11824 |
| <i>MultiMed-ST: Large-scale Many-to-many Multilingual Medical Speech Translation</i>                                                                                                                              |       |
| Khai Le-Duc, Tuyen Tran, Bach Phan Tat, Nguyen Kim Hai Bui, Quan Dang Anh, Hung-Phong Tran, Thanh Thuy Nguyen, Ly Nguyen, Tuan Minh Phan, Thi Thu Phuong Tran, Chris Ngo, Khanh Xuan Nguyen and Thanh Nguyen-Tang | 11839 |
| <i>Beyond Checkmate: Exploring the Creative Choke Points for AI Generated Texts</i>                                                                                                                               |       |
| Nafis Irtiza Tripto, Saranya Venkatraman, Mahjabin Nahar and Dongwon Lee                                                                                                                                          | 11965 |
| <i>MoR: Better Handling Diverse Queries with a Mixture of Sparse, Dense, and Human Retrievers</i>                                                                                                                 |       |
| Jushaan Singh Kalra, Xinran Zhao, To Eun Kim, Fengyu Cai, Fernando Diaz and Tongshuang Wu                                                                                                                         | 11983 |
| <i>Learning Contextual Retrieval for Robust Conversational Search</i>                                                                                                                                             |       |
| Seunghan Yang, Juntae Lee, Jihwan Bang, Kyuhong Shim, Minsoo Kim and Simyung Chang                                                                                                                                | 12003 |
| <i>LIDDIA: Language-based Intelligent Drug Discovery Agent</i>                                                                                                                                                    |       |
| Reza Averly, Frazier N. Baker, Ian A Watson and Xia Ning                                                                                                                                                          | 12016 |
| <i>Agentic-R1: Distilled Dual-Strategy Reasoning</i>                                                                                                                                                              |       |
| Weihua Du, Pranjal Aggarwal, Sean Welleck and Yiming Yang                                                                                                                                                         | 12041 |
| <i>Proactive Assistant Dialogue Generation from Streaming Egocentric Videos</i>                                                                                                                                   |       |
| Yichi Zhang, Xin Luna Dong, Zhaojiang Lin, Andrea Madotto, Anuj Kumar, Babak Damavandi, Joyce Chai and Seungwhan Moon                                                                                             | 12056 |
| <i>Should I Share this Translation? Evaluating Quality Feedback for User Reliance on Machine Translation</i>                                                                                                      |       |
| Dayeon Ki, Kevin Duh and Marine Carpuat                                                                                                                                                                           | 12081 |
| <i>ChartGaze: Enhancing Chart Understanding in LVLMs with Eye-Tracking Guided Attention Refinement</i>                                                                                                            |       |
| Ali Salamatian, Amirhossein Abaskohi, Wan-Cyuan Fan, Mir Rayat Imtiaz Hossain, Leonid Sigal and Giuseppe Carenini                                                                                                 | 12105 |
| <i>LogiCoL: Logically-Informed Contrastive Learning for Set-based Dense Retrieval</i>                                                                                                                             |       |
| Yanzhen Shen, Sihao Chen, Xueqiang Xu, Yunyi Zhang, Chaitanya Malaviya and Dan Roth                                                                                                                               | 12126 |

|                                                                                                                                |       |
|--------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>ModalPrompt: Towards Efficient Multimodal Continual Instruction Tuning with Dual-Modality Guided Prompt</i>                 |       |
| Fanhu Zeng, Fei Zhu, Haiyang Guo, Xu-Yao Zhang and Cheng-Lin Liu . . . . .                                                     | 12138 |
| <i>Skip-Thinking: Chunk-wise Chain-of-Thought Distillation Enable Smaller Language Models to Reason Better and Faster</i>      |       |
| Xiaoshu Chen, Sihang Zhou, KE Liang, Xiaoyu Sun and Xinwang Liu . . . . .                                                      | 12154 |
| <i>Can an Individual Manipulate the Collective Decisions of Multi-Agents?</i>                                                  |       |
| Fengyuan Liu, Rui Zhao, Shuo Chen, Guohao Li, Philip Torr, Lei Han and Jindong Gu . . .                                        | 12170 |
| <i>Toxicity Red-Teaming: Benchmarking LLM Safety in Singapore’s Low-Resource Languages</i>                                     |       |
| Yujia Hu, Ming Shan Hee, Preslav Nakov and Roy Ka-Wei Lee . . . . .                                                            | 12195 |
| <i>Improving Clustering with Positive Pairs Generated from LLM-Driven Labels</i>                                               |       |
| Xiaotong Zhang and Ying Li . . . . .                                                                                           | 12214 |
| <i>Gamma-Guard: Lightweight Residual Adapters for Robust Guardrails in Large Language Models</i>                               |       |
| Lijia Lv, Yuanshu Zhao, Guan Wang, Xuehai Tang, Wen Jie, Jizhong Han and Songlin Hu                                            | 12231 |
| <i>Facilitating Long Context Understanding via Supervised Chain-of-Thought Reasoning</i>                                       |       |
| Jingyang Lin, Andy Wong, Tian Xia, Shenghua He, Hui Wei, Mei Han and Jiebo Luo . . .                                           | 12244 |
| <i>Dynamic Energy-Based Contrastive Learning with Multi-Stage Knowledge Verification for Event Causality Identification</i>    |       |
| Ya Su, Hu Zhang, Yue Fan, Guangjun Zhang, YuJie Wang, Ru Li and Hongye Tan . . . . .                                           | 12261 |
| <i>ICG: Improving Cover Image Generation via MLLM-based Prompting and Personalized Preference Alignment</i>                    |       |
| Zhipeng Bian, Jieming Zhu, Qijiong Liu, Wang Lin, Guohao Cai, Zhaocheng Du, Jiacheng Sun, Zhou Zhao and Zhenhua Dong . . . . . | 12280 |
| <i>From Long to Lean: Performance-aware and Adaptive Chain-of-Thought Compression via Multi-round Refinement</i>               |       |
| JianZhi Yan, Le Liu, Youcheng Pan, Shiwei Chen, Zike Yuan, Yang Xiang and Buzhou Tang                                          | 12291 |
| <i>A Symbolic Adversarial Learning Framework for Evolving Fake News Generation and Detection</i>                               |       |
| Chong Tian, Qirong Ho and Xiuying Chen . . . . .                                                                               | 12308 |
| <i>RareSyn: Health Record Synthesis for Rare Disease Diagnosis</i>                                                             |       |
| Huimin Wang, Yutian Zhao, Yefeng Zheng and Xian Wu . . . . .                                                                   | 12323 |
| <i>Sticker-TTS: Learn to Utilize Historical Experience with a Sticker-driven Test-Time Scaling Framework</i>                   |       |
| Jie Chen, Jinhao Jiang, Yingqian Min, Zican Dong, Shijie Wang, Xin Zhao and Ji-Rong Wen                                        | 12340 |
| <i>CMHG: A Dataset and Benchmark for Headline Generation of Minority Languages in China</i>                                    |       |
| Guixian Xu, Zeli Su, Ziyin Zhang, Jianing Liu, Xu Han, Ting Zhang and Yushuang Dong .                                          | 12351 |
| <i>Understanding the Information Propagation Effects of Communication Topologies in LLM-based Multi-Agent Systems</i>          |       |
| Xu Shen, Yixin Liu, Yiwei Dai, Yili Wang, Rui Miao, Yue Tan, Shirui Pan and Xin Wang .                                         | 12359 |
| <i>Boosting Data Utilization for Multilingual Dense Retrieval</i>                                                              |       |
| Chao Huang, Fengran Mo, Yufeng Chen, Changhao Guan, Zhenrui Yue, Xinyu Wang, Jinan Xu and Kaiyu Huang . . . . .                | 12374 |

|                                                                                                                                       |       |
|---------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Self-Augmented Preference Alignment for Sycophancy Reduction in LLMs</i>                                                           |       |
| Chien Hung Chen, Hen-Hsen Huang and Hsin-Hsi Chen .....                                                                               | 12391 |
| <i>TP-RAG: Benchmarking Retrieval-Augmented Large Language Model Agents for Spatiotemporal-Aware Travel Planning</i>                  |       |
| Hang Ni, Fan Liu, Xinyu Ma, Lixin Su, Shuaiqiang Wang, Dawei Yin, Hui Xiong and Hao Liu                                               | 12404 |
| <i>Recontextualizing Revitalization: A Mixed Media Approach to Reviving the Nüshu Language</i>                                        |       |
| Ivory Yang, Xiaobo Guo, Yuxin Wang, Hefan Zhang, Yaning Jia, William Dinauer and Soroush Vosoughi .....                               | 12431 |
| <i>Towards Advanced Mathematical Reasoning for LLMs via First-Order Logic Theorem Proving</i>                                         |       |
| Chuxue Cao, Mengze Li, Juntao Dai, Jinluan Yang, Zijian Zhao, Shengyu Zhang, Weijie Shi, Chengzhong Liu, Sirui Han and Yike Guo ..... | 12441 |
| <i>From Tens of Hours to Tens of Thousands: Scaling Back-Translation for Speech Recognition</i>                                       |       |
| Tianduo Wang, Lu Xu, Wei Lu and Shanbo Cheng .....                                                                                    | 12462 |
| <i>CityEQA: A Hierarchical LLM Agent on Embodied Question Answering Benchmark in City Space</i>                                       |       |
| Yong Zhao, Kai Xu, Zhengqiu Zhu, Yue Hu, Zhiheng Zheng, Yingfeng Chen, Yatai Ji, Chen Gao, Yong Li and Jincai Huang .....             | 12477 |
| <i>Mitigating Hallucinations in Vision-Language Models through Image-Guided Head Suppression</i>                                      |       |
| Sreetama Sarkar, Yue Che, Alex Gavin, Peter Anthony Beerel and Souvik Kundu .....                                                     | 12493 |
| <i>Examining False Positives under Inference Scaling for Mathematical Reasoning</i>                                                   |       |
| Yu Wang, Nan Yang, Liang Wang, Furu Wei and Fuli Feng .....                                                                           | 12513 |
| <i>Translationese-index: Using Likelihood Ratios for Graded and Generalizable Measurement of Translationese</i>                       |       |
| Yikang Liu, Wanyang Zhang, Yiming Wang, Jialong Tang, Pei Zhang, Baosong Yang, Fei Huang, Rui Wang and Hai Hu .....                   | 12533 |
| <i>Exploring the Limitations of Mamba in COPY and CoT Reasoning</i>                                                                   |       |
| Ruifeng Ren, Zhicong Li and Yong Liu .....                                                                                            | 12551 |
| <i>ProcWorld: Benchmarking Large Model Planning in Reachability-Constrained Environments</i>                                          |       |
| Dong Wang, Xinghang Li, Zhengshen Zhang, Jirong Liu, Xiao Ma, Hanbo Zhang, Tao Kong and Huaping Liu .....                             | 12576 |
| <i>R2I-Bench: Benchmarking Reasoning-Driven Text-to-Image Generation</i>                                                              |       |
| Kaijie Chen, Zihao Lin, Zhiyang Xu, Ying Shen, Yuguang Yao, Joy Rimchala, Jiaxin Zhang and Lifu Huang .....                           | 12607 |
| <i>Can GRPO Boost Complex Multimodal Table Understanding?</i>                                                                         |       |
| Xiaoqiang Kang, Shengen Wu, Zimu Wang, Yilin Liu, Xiaobo Jin, Kaizhu Huang, Wei Wang, Yutao Yue, Xiaowei Huang and Qiufeng Wang ..... | 12643 |
| <i>MoMoE: Mixture of Moderation Experts Framework for AI-Assisted Online Governance</i>                                               |       |
| Agam Goyal, Xianyang Zhan, Yilun Chen, Koustuv Saha and Eshwar Chandrasekharan ...                                                    | 12657 |
| <i>Following the Autoregressive Nature of LLM Embeddings via Compression and Alignment</i>                                            |       |
| Jingcheng Deng, Zhongtao Jiang, Liang Pang, Zihao Wei, Liwei Chen, Kun Xu, Yang Song, Huawei Shen and Xueqi Cheng .....               | 12673 |

|                                                                                                                                                |       |
|------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Evaluating LLM-Generated Diagrams as Graphs</i>                                                                                             |       |
| Chumeng Liang and Jiaxuan You . . . . .                                                                                                        | 12690 |
| <i>Breaking Bad Tokens: Detoxification of LLMs Using Sparse Autoencoders</i>                                                                   |       |
| Agam Goyal, Vedant Rathi, William Yeh, Yian Wang, Yuen Chen and Hari Sundaram . . . .                                                          | 12703 |
| <i>VCSearch: Bridging the Gap Between Well-Defined and Ill-Defined Problems in Mathematical Reasoning</i>                                      |       |
| Shi-Yu Tian, Zhi Zhou, Kun-Yang Yu, Ming Yang, Lin-Han Jia, Lan-Zhe Guo and Yu-Feng Li                                                         | 12722 |
| <i>How do autoregressive transformers solve full addition?</i>                                                                                 |       |
| Wang Peixu, Chen Yu, Yu Ming and Cheng Xiang . . . . .                                                                                         | 12744 |
| <i>MAIN: Mutual Alignment Is Necessary for instruction tuning</i>                                                                              |       |
| Fanyi Yang, Jianfeng Liu, Xin Zhang, Haoyu Liu, Xixin Cao, Yuefeng Zhan, Hao Sun, Weiwei Deng, Feng Sun and Qi Zhang . . . . .                 | 12769 |
| <i>Expanding before Inferring: Enhancing Factuality in Large Language Models through Premature Layers Interpolation</i>                        |       |
| Dingwei Chen, Ziqiang Liu, Feiteng Fang, Chak Tou Leong, Shiwen Ni, Ahmadreza Argha, Hamid Alinejad-Rokny, Min Yang and Chengming Li . . . . . | 12782 |
| <i>DeepWell-Adol: A Scalable Expert-Based Dialogue Corpus for Adolescent Positive Mental Health and Wellbeing Promotion</i>                    |       |
| Wenyu Qiu, Yuxiong Wang, Jiajun Tan, Hanchao Hou, Qinda Liu, Wei Yao and Shiguang NI                                                           | 12798 |
| <i>Data to Defense: The Role of Curation in Aligning Large Language Models Against Safety Compromise</i>                                       |       |
| Xiaoqun Liu, Jiacheng Liang, Luoxi Tang, Muchao Ye, Weicheng Ma and Zhaoan Xi . . . .                                                          | 12823 |
| <i>Speculative Safety-Aware Decoding</i>                                                                                                       |       |
| Xuekang Wang, Shengyu Zhu and Xueqi Cheng . . . . .                                                                                            | 12839 |
| <i>PanicToCalm: A Proactive Counseling Agent for Panic Attacks</i>                                                                             |       |
| Jihyun Lee, Yejin Min, San Kim, Yejin Jeon, Sung Jun Yang, Hyounghun Kim and Gary Lee                                                          | 12854 |
| <i>CoPL: Collaborative Preference Learning for Personalizing LLMs</i>                                                                          |       |
| Youngbin Choi, Seunghyuk Cho, Minjong Lee, MoonJeong Park, Yesong Ko, Jungseul Ok and Dongwoo Kim . . . . .                                    | 12887 |
| <i>Dynamic Collaboration of Multi-Language Models based on Minimal Complete Semantic Units</i>                                                 |       |
| Chao Hao, Zezheng Wang, Yanhua Huang, Ruiwen Xu, Wenzhe Niu, Xin Liu and Zitong YU                                                             | 12906 |
| <i>AI Chatbots as Professional Service Agents: Developing a Professional Identity</i>                                                          |       |
| Wenwen Li, Kangwei Shi and Yidong Chai . . . . .                                                                                               | 12924 |
| <i>DeepResonance: Enhancing Multimodal Music Understanding via Music-centric Multi-way Instruction Tuning</i>                                  |       |
| Zhuoyuan Mao, Mengjie Zhao, Qiyu Wu, Hiromi Wakaki and Yuki Mitsufuji . . . . .                                                                | 12938 |
| <i>Advancing Oversight Reasoning across Languages for Audit Sycophantic Behaviour via X-Agent</i>                                              |       |
| Giulia Pucci and Leonardo Ranaldi . . . . .                                                                                                    | 12961 |

|                                                                                                                                                                                                |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>CAFE: Retrieval Head-based Coarse-to-Fine Information Seeking to Enhance Multi-Document QA Capability</i>                                                                                   |       |
| Han Peng, Jinhao Jiang, Zican Dong, Xin Zhao and Lei Fang . . . . .                                                                                                                            | 12978 |
| <i>SSA-COMET: Do LLMs Outperform Learned Metrics in Evaluating MT for Under-Resourced African Languages?</i>                                                                                   |       |
| Senyu Li, Jiayi Wang, Felermينو D. M. A. Ali, Colin Cherry, Daniel Deutsch, Eleftheria Briakou, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp and David Ifeoluwa Adelani . . . . . | 12991 |
| <i>FaithUn: Toward Faithful Forgetting in Language Models by Investigating the Interconnectedness of Knowledge</i>                                                                             |       |
| Nakyeong Yang, Minsung Kim, Seunghyun Yoon, Joongbo Shin and Kyomin Jung . . . . .                                                                                                             | 13011 |
| <i>Calibrating Pseudo-Labeling with Class Distribution for Semi-supervised Text Classification</i>                                                                                             |       |
| Weiyi Yang, Richong Zhang, Junfan Chen and Jiawei Sheng . . . . .                                                                                                                              | 13027 |
| <i>Coarse-to-Fine Grounded Memory for LLM Agent Planning</i>                                                                                                                                   |       |
| Wei Yang, Jinwei Xiao, Hongming Zhang, Qingyang Zhang, Yanna Wang and bo xu . . . . .                                                                                                          | 13041 |
| <i>From A and B to A+B: Can Large Language Models Solve Compositional Math Problems?</i>                                                                                                       |       |
| Xisheng Xiao and Hanlin Zhao . . . . .                                                                                                                                                         | 13069 |
| <i>Sycophancy Mitigation Through Reinforcement Learning with Uncertainty-Aware Adaptive Reasoning Trajectories</i>                                                                             |       |
| Mohammad Beigi, Ying Shen, Parshin Shojaee, Qifan Wang, Zichao Wang, Chandan K. Reddy, Ming Jin and Lifu Huang . . . . .                                                                       | 13091 |
| <i>SimVBG: Simulating Individual Values by Backstory Generation</i>                                                                                                                            |       |
| Bangde Du, Ziyi Ye, Zhijing Wu, Monika A. Jankowska, Shuqi Zhu, Qingyao Ai, Yujia Zhou and Yiqun Liu . . . . .                                                                                 | 13105 |
| <i>EvolveSearch: An Iterative Self-Evolving Search Agent</i>                                                                                                                                   |       |
| Ding-Chu Zhang, Yida Zhao, Jialong Wu, Liwen Zhang, Baixuan Li, Wenbiao Yin, Yong Jiang, Yu-Feng Li, Kewei Tu, Pengjun Xie and Fei Huang . . . . .                                             | 13135 |
| <i>Syntax-Aware Retrieval Augmentation for Neural Symbolic Regression</i>                                                                                                                      |       |
| Canmiao Zhou and Han Huang . . . . .                                                                                                                                                           | 13149 |
| <i>Merge then Realign: Simple and Effective Modality-Incremental Continual Learning for Multimodal LLMs</i>                                                                                    |       |
| Dingkun Zhang, Shuhan Qi, Xinyu Xiao, Kehai Chen and Xuan Wang . . . . .                                                                                                                       | 13160 |
| <i>Graceful Forgetting in Generative Language Models</i>                                                                                                                                       |       |
| Chunyang Jiang, Chi-Min Chan, Yiyang Cai, Yulong Liu, Wei Xue and Yike Guo . . . . .                                                                                                           | 13177 |
| <i>Answering Narrative-Driven Recommendation Queries via a Retrieve-Rank Paradigm and the OCG-Agent</i>                                                                                        |       |
| Yunxiao Shi, Haoning Shang, Xing Zi, Wujiang Xu, Yue Feng and Min Xu . . . . .                                                                                                                 | 13193 |
| <i>Direct Value Optimization: Improving Chain-of-Thought Reasoning in LLMs with Refined Values</i>                                                                                             |       |
| Hongbo Zhang, Han Cui, Guangsheng Bao, Linyi Yang, Jun Wang and Yue Zhang . . . . .                                                                                                            | 13215 |
| <i>Jailbreak-Tuning: Models Efficiently Learn Jailbreak Susceptibility</i>                                                                                                                     |       |
| Brendan Murphy, Dillon Bowen, Shahrad Mohammadzadeh, Tom Tseng, Julius Broomfield, Adam Gleave and Kellin Pelrine . . . . .                                                                    | 13229 |

|                                                                                                                                                 |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Neural Topic Modeling via Contextual and Graph Information Fusion</i>                                                                        |       |
| Jiyuan Liu, Jiaying Yan, Chunjiang Zhu, Xingyu Liu, Li Qing and Yanghui Rao . . . . .                                                           | 13259 |
| <i>CARE: A Disagreement Detection Framework with Concept Alignment and Reasoning Enhancement</i>                                                |       |
| Jiyuan Liu, Jielin Song, Yunhe Pang, Zhiyu Shen and Yanghui Rao . . . . .                                                                       | 13276 |
| <i>Beyond Task-Oriented and Chitchat Dialogues: Proactive and Transition-Aware Conversational Agents</i>                                        |       |
| Yejin Yoon, Yuri Son, Namyong So, Minseo Kim, Minsoo Cho, Chanhee Park, Seungshin Lee and Taeuk Kim . . . . .                                   | 13292 |
| <i>LightThinker: Thinking Step-by-Step Compression</i>                                                                                          |       |
| Jintian Zhang, Yuqi Zhu, Mengshu Sun, Yujie Luo, Shuofei Qiao, Lun Du, Da Zheng, Huajun Chen and Ningyu Zhang . . . . .                         | 13319 |
| <i>How Is LLM Reasoning Distracted by Irrelevant Context? An Analysis Using a Controlled Benchmark</i>                                          |       |
| Minglai Yang, Ethan Huang, Liang Zhang, Mihai Surdeanu, William Yang Wang and Liangming Pan . . . . .                                           | 13341 |
| <i>Investigating Pedagogical Teacher and Student LLM Agents: Genetic Adaptation Meets Retrieval-Augmented Generation Across Learning Styles</i> |       |
| Debdeep Sanyal, Agniva Maiti, Umakanta Maharana, Dhruv Kumar, Ankur Mali, C. Lee Giles and Murari Mandal . . . . .                              | 13360 |
| <i>GeoEdit: Geometric Knowledge Editing for Large Language Models</i>                                                                           |       |
| Yujie Feng, Li-Ming Zhan, Zexin LU, Yongxin Xu, Xu Chu, Yasha Wang, Jiannong Cao, Philip S. Yu and Xiao-Ming Wu . . . . .                       | 13402 |
| <i>A Generative Pre-Trained Language Model for Channel Prediction in Wireless Communications Systems</i>                                        |       |
| Bo Lin, Huanming Zhang, Yuhua Jiang, Yucong Wang, Tengyu Zhang, Shaoqiang Yan, Hongyao Li, Yihong Liu and Feifei Gao . . . . .                  | 13418 |
| <i>AIMMerging: Adaptive Iterative Model Merging Using Training Trajectories for Language Model Continual Learning</i>                           |       |
| Yujie Feng, Jian Li, Xiaoyu Dong, Pengfei Xu, Xiaohui Zhou, Yujia Zhang, Zexin LU, Yasha Wang, Alan Zhao, Xu Chu and Xiao-Ming Wu . . . . .     | 13432 |
| <i>R-PRM: Reasoning-Driven Process Reward Modeling</i>                                                                                          |       |
| Shuaijie She, Junxiao Liu, Yifeng Liu, Jiajun Chen, Xin Huang and Shujian Huang . . . . .                                                       | 13450 |
| <i>RLAE: Reinforcement Learning-Assisted Ensemble for LLMs</i>                                                                                  |       |
| Yuqian Fu, Yuanheng Zhu, Jiajun Chai, Guojun Yin, Wei Lin, Qichao Zhang and Dongbin Zhao                                                        | 13464 |
| <i>Do Large Language Models Truly Grasp Addition? A Rule-Focused Diagnostic Using Two-Integer Arithmetic</i>                                    |       |
| Yang Yan, Yu Lu, Renjun Xu and Zhenzhong Lan . . . . .                                                                                          | 13479 |
| <i>AskToAct: Enhancing LLMs Tool Use via Self-Correcting Clarification</i>                                                                      |       |
| Xuan Zhang, Yongliang Shen, Zhe Zheng, Linjuan Wu, Wenqi Zhang, Yuchen Yan, Qiuying Peng, Jun Wang and Weiming Lu . . . . .                     | 13496 |
| <i>START: Self-taught Reasoner with Tools</i>                                                                                                   |       |
| Chengpeng Li, Mingfeng Xue, Zhenru Zhang, Jiayi Yang, Beichen Zhang, Bowen Yu, Binyuan Hui, Junyang Lin, Xiang Wang and Dayiheng Liu . . . . .  | 13524 |

|                                                                                                                                                                                           |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>The Impact of Negated Text on Hallucination with Large Language Models</i>                                                                                                             |       |
| Jaehyung Seo, Hyeonseok Moon and Heuseok Lim .....                                                                                                                                        | 13566 |
| <i>A Probabilistic Inference Scaling Theory for LLM Self-Correction</i>                                                                                                                   |       |
| Zhe Yang, Yichang Zhang, Yudong Wang, Ziyao Xu, Junyang Lin and Zhifang Sui .....                                                                                                         | 13585 |
| <i>MentalGLM Series: Explainable Large Language Models for Mental Health Analysis on Chinese Social Media</i>                                                                             |       |
| Wei Zhai, Nan Bai, Qing Zhao, Jianqiang Li, Fan Wang, Hongzhi Qi, Meng Jiang, Xiaoqin Wang, Bing Xiang Yang and Guanghui FU .....                                                         | 13600 |
| <i>Knowledge-Aware Co-Reasoning for Multidisciplinary Collaboration</i>                                                                                                                   |       |
| Xurui li, Wanghailiao, Kaisong Song, Rui Zhu and Haixu Tang .....                                                                                                                         | 13616 |
| <i>Astra: Efficient Transformer Architecture and Contrastive Dynamics Learning for Embodied Instruction Following</i>                                                                     |       |
| Yueen Ma, DaFeng Chi, Shiguang Wu, Yuecheng Liu, Yuzheng Zhuang and Irwin King ..                                                                                                         | 13633 |
| <i>MAVL: A Multilingual Audio-Video Lyrics Dataset for Animated Song Translation</i>                                                                                                      |       |
| Woohyun Cho, Youngmin Kim, Sunghyun Lee and Youngjae Yu .....                                                                                                                             | 13652 |
| <i>MuTIS: Enhancing Reasoning Efficiency through Multi Turn Intervention Sampling in Reinforcement Learning</i>                                                                           |       |
| Wenshuo Zhao, Haoxing Zhai, Xinyu Qiu, Zhenting Qi, Shuhe Li and Linchao Zhu .....                                                                                                        | 13681 |
| <i>PRIM: Towards Practical In-Image Multilingual Machine Translation</i>                                                                                                                  |       |
| Yanzhi Tian, Zeming Liu, Zhengyang Liu, Chong Feng, Xin Li, Heyan Huang and Yuhang Guo                                                                                                    | 13694 |
| <i>Mind the Inclusivity Gap: Multilingual Gender-Neutral Translation Evaluation with mGeNTE</i>                                                                                           |       |
| Beatrice Savoldi, Giuseppe Attanasio, Eleonora Cupin, Eleni Gkovedarou, Janiça Hackenbuchner, Anne Lauscher, Matteo Negri, Andrea Piergentili, Manjinder Thind and Luisa Bentivogli ..... | 13710 |
| <i>DiplomacyAgent: Do LLMs Balance Interests and Ethical Principles in International Events?</i>                                                                                          |       |
| Jianxiang Peng, Ling Shi, Xinwei Wu, Hanwen Zhang, Fujian Liu, Haocheng Lyu and Deyi Xiong .....                                                                                          | 13733 |
| <i>DisLoRA: Task-specific Low-Rank Adaptation via Orthogonal Basis from Singular Value Decomposition</i>                                                                                  |       |
| She Yifei, Xinhao Wei and Yulong Wang .....                                                                                                                                               | 13752 |
| <i>Unmasking Deceptive Visuals: Benchmarking Multimodal Large Language Models on Misleading Chart Question Answering</i>                                                                  |       |
| Zixin Chen, Sicheng Song, KaShun Shum, Yanna Lin, Rui Sheng, Weiqi Wang and Huamin Qu                                                                                                     | 13768 |
| <i>Textual Aesthetics in Large Language Models</i>                                                                                                                                        |       |
| Lingjie Jiang, Shaohan Huang, Xun Wu and Furu Wei .....                                                                                                                                   | 13802 |
| <i>Section-Level Simplification of Biomedical Abstracts</i>                                                                                                                               |       |
| Jan Bakker and Jaap Kamps .....                                                                                                                                                           | 13831 |
| <i>PoseStitch-SLT: Linguistically Inspired Pose-Stitching for End-to-End Sign Language Translation</i>                                                                                    |       |
| Abhinav Joshi, Vaibhav Sharma, Sanjeet Singh and Ashutosh Modi .....                                                                                                                      | 13846 |
| <i>Few-Shot Open-Set Classification via Reasoning-Aware Decomposition</i>                                                                                                                 |       |
| Avyav Kumar Singh and Helen Yannakoudakis .....                                                                                                                                           | 13866 |



|                                                                                                                                                |       |
|------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Translation in the Hands of Many: Centering Lay Users in Machine Translation Interactions</i>                                               |       |
| Beatrice Savoldi, Alan Ramponi, Matteo Negri and Luisa Bentivogli . . . . .                                                                    | 13888 |
| <i>iTool: Reinforced Fine-Tuning with Dynamic Deficiency Calibration for Advanced Tool Use</i>                                                 |       |
| Yirong Zeng, Xiao Ding, Yuxian Wang, Weiwen Liu, Yutai Hou, Wu Ning, Xu Huang, Duyu Tang, Dandan Tu, Bing Qin and Ting Liu . . . . .           | 13902 |
| <i>Transplant Then Regenerate: A New Paradigm for Text Data Augmentation</i>                                                                   |       |
| Guangzhan Wang, Hongyu Zhang, Beijun Shen and Xiaodong Gu . . . . .                                                                            | 13918 |
| <i>Compositional Generalisation for Explainable Hate Speech Detection</i>                                                                      |       |
| Agostina Calabrese, Tom Sherborne, Björn Ross and Mirella Lapata . . . . .                                                                     | 13933 |
| <i>CCQA: Generating Question from Solution Can Improve Inference-Time Reasoning in SLMs</i>                                                    |       |
| Jinyoung Kim and Ji Won Yoon . . . . .                                                                                                         | 13956 |
| <i>TVQACML: Benchmarking Text-Centric Visual Question Answering in Multilingual Chinese Minority Languages</i>                                 |       |
| Sha Jiu, Yu Weng, Mengxiao Zhu, Chong Feng, Zheng Liu and Jialedongzhu . . . . .                                                               | 13969 |
| <i>Transparent and Coherent Procedural Mistake Detection</i>                                                                                   |       |
| Shane Storks, Itamar Bar-Yossef, Yayuan Li, Zheyuan Zhang, Jason J Corso and Joyce Chai                                                        | 13980 |
| <i>Teaching Your Models to Understand Code via Focal Preference Alignment</i>                                                                  |       |
| Jie Wu, Haoling Li, Xin Zhang, Xiao Liu, Yangyu Huang, Jianwen Luo, Yizhen Zhang, Zuchao Li, Ruihang Chu, Yujiu Yang and Scarlett Li . . . . . | 14015 |
| <i>MoLoRAG: Bootstrapping Document Understanding via Multi-modal Logic-aware Retrieval</i>                                                     |       |
| Xixi Wu, Yanchao Tan, Nan Hou, Ruiyang Zhang and Hong Cheng . . . . .                                                                          | 14036 |
| <i>Vision-Free Retrieval: Rethinking Multimodal Search with Textual Scene Descriptions</i>                                                     |       |
| Ioanna Ntinou, Alexandros Xenos, Yassine Ouali, Adrian Bulat and Georgios Tzimiropoulos                                                        | 14058 |
| <i>TableRAG: A Retrieval Augmented Generation Framework for Heterogeneous Document Reasoning</i>                                               |       |
| Xiaohan Yu, Pu Jian and Chong Chen . . . . .                                                                                                   | 14075 |
| <i>Retrieval Enhanced Feedback via In-context Neural Error-book</i>                                                                            |       |
| Jongyeop Hyun and Bumsu Kim . . . . .                                                                                                          | 14095 |
| <i>Improve LLM-as-a-Judge Ability as a General Ability</i>                                                                                     |       |
| Jiachen Yu, Shaoning Sun, Xiaohui Hu, Jiaxu Yan, Kaidong Yu and Xuelong Li . . . . .                                                           | 14111 |
| <i>G2: Guided Generation for Enhanced Output Diversity in LLMs</i>                                                                             |       |
| Zhiwen Ruan, Yixia Li, Yefeng Liu, Yun Chen, Weihua Luo, Peng Li, Yang Liu and Guanhua Chen . . . . .                                          | 14128 |
| <i>ToolSafety: A Comprehensive Dataset for Enhancing Safety in LLM-Based Agent Tool Invocations</i>                                            |       |
| Yuejin Xie, Youliang Yuan, Wenxuan Wang, Fan Mo, Jianmin Guo and Pinjia He . . . . .                                                           | 14147 |
| <i>Learning to See through Sound: From VggCaps to Multi2Cap for Richer Automated Audio Captioning</i>                                          |       |
| Sangyeon Cho, Mingi Kim, Jinkwon Hwang, Jaehoon Go, Minuk Ma, Sunjae Yoon and Junyeong Kim . . . . .                                           | 14169 |
| <i>Towards Optimal Evaluation Efficiency for Large Language Models</i>                                                                         |       |
| Guohong Li and Deyi Xiong . . . . .                                                                                                            | 14188 |



|                                                                                                                                       |       |
|---------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>MMAPG: A Training-Free Framework for Multimodal Multi-hop Question Answering via Adaptive Planning Graphs</i>                      |       |
| Yiheng Hu, Xiaoyang Wang, Qing Liu, Xiwei Xu, Qian Fu, Wenjie Zhang and Liming Zhu                                                    | 14196 |
| <i>Mixture-of-Clustered-Experts: Advancing Expert Specialization and Generalization in Instruction Tuning</i>                         |       |
| Sugyeong Eo, Jung Jun Lee, Chanjun Park and Heuiseok Lim                                                                              | 14213 |
| <i>Process-Supervised Reinforcement Learning for Code Generation</i>                                                                  |       |
| Yufan Ye, Ting Zhang, Wenbin Jiang and Hua Huang                                                                                      | 14225 |
| <i>MuCAL: Contrastive Alignment for Preference-Driven KG-to-Text Generation</i>                                                       |       |
| Yifei Song and Claire Gardent                                                                                                         | 14239 |
| <i>Advancing Fine-Grained Visual Understanding with Multi-Scale Alignment in Multi-Modal Models</i>                                   |       |
| Wei Wang, Zhaowei Li, Qi Xu, Linfeng Li, YiQing Cai, Botian Jiang, Hang Song, Xingcan Hu, Pengyu Wang and Li Xiao                     | 14283 |
| <i>Thought calibration: Efficient and confident test-time scaling</i>                                                                 |       |
| Menghua Wu, Cai Zhou, Stephen Bates and Tommi Jaakkola                                                                                | 14303 |
| <i>Can LLMs Reason Abstractly Over Math Word Problems Without CoT? Disentangling Abstract Formulation From Arithmetic Computation</i> |       |
| Ziling Cheng, Meng Cao, Leila Pishdad, Yanshuai Cao and Jackie CK Cheung                                                              | 14318 |
| <i>QCRD: Quality-guided Contrastive Rationale Distillation for Large Language Models</i>                                              |       |
| Wei Wang, Zhaowei Li, Qi Xu, YiQing Cai, Hang Song, Qi Qi, Ran Zhou, Zhida Huang, Tao Wang and Li Xiao                                | 14346 |
| <i>SHARP: Steering Hallucination in LVLMs via Representation Engineering</i>                                                          |       |
| Junfei Wu, Yue Ding, Guofan Liu, Tianze Xia, Ziyue Huang, Dianbo Sui, Qiang Liu, Shu Wu, Liang Wang and Tieniu Tan                    | 14358 |
| <i>Think, Verbalize, then Speak: Bridging Complex Thoughts and Comprehensible Speech</i>                                              |       |
| Tony Woo, Sehun Lee, Kang-wook Kim and Gunhee Kim                                                                                     | 14374 |
| <i>Warm Up Before You Train: Unlocking General Reasoning in Resource-Constrained Settings</i>                                         |       |
| Safal Shrestha, Minwu Kim, Aadim Nepal, Anubhav Shrestha and Keith W. Ross                                                            | 14392 |
| <i>PPTAgent: Generating and Evaluating Presentations Beyond Text-to-Slides</i>                                                        |       |
| Hao Zheng, Xinyan Guan, Hao Kong, Wenkai Zhang, Jia Zheng, Weixiang Zhou, Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun               | 14414 |
| <i>SWAM: Adaptive Sliding Window and Memory-Augmented Attention Model for Rumor Detection</i>                                         |       |
| Mei Guo, Chen Chen, Chunyan Hou, Yike Wu and Xiaojie Yuan                                                                             | 14431 |
| <i>HydraRAG: Structured Cross-Source Enhanced Large Language Model Reasoning</i>                                                      |       |
| Xingyu Tan, Xiaoyang Wang, Qing Liu, Xiwei Xu, Xin Yuan, Liming Zhu and Wenjie Zhang                                                  | 14443 |
| <i>VRoPE: Rotary Position Embedding for Video Large Language Models</i>                                                               |       |
| Zikang Liu, Longteng Guo, Yepeng Tang, Tongtian Yue, Junxian Cai, Kai Ma, Qingbin Liu, Xi Chen and Jing Liu                           | 14472 |
| <i>SciNLP: A Domain-Specific Benchmark for Full-Text Scientific Entity and Relation Extraction in NLP</i>                             |       |
| Decheng Duan, Jitong Peng, Yingyi Zhang and Chengzhi Zhang                                                                            | 14485 |

|                                                                                                                                                 |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Think and Recall: Layer-Level Prompting for Lifelong Model Editing</i>                                                                       |       |
| Jinke Wang, Zenan Ying, Qi Liu, Wei Chen, Tong Xu, Huijun Hou and Zhi Zheng . . . . .                                                           | 14499 |
| <i>SPIRIT: Patching Speech Language Models against Jailbreak Attacks</i>                                                                        |       |
| Amirbek Djanibekov, Nurdaulet Mukhituly, Kentaro Inui, Hanan Aldarmaki and Nils Lukas                                                           | 14515 |
| <i>FIRE: Flexible Integration of Data Quality Ratings for Effective Pretraining</i>                                                             |       |
| Xu Liangyu, Xuemiao Zhang, Feiyu Duan, Sirui Wang, Rongxiang Weng, Jingang Wang and Xunliang Cai . . . . .                                      | 14533 |
| <i>Multi-Domain Explainability of Preferences</i>                                                                                               |       |
| Nitay Calderon, Liat Ein-Dor and Roi Reichart . . . . .                                                                                         | 14554 |
| <i>Tuning Less, Prompting More: In-Context Preference Learning Pipeline for Natural Language Transformation</i>                                 |       |
| Shuyun Yang, Yan Zhang, Zhengmao Ye, Lei Duan and Mingjie Tang . . . . .                                                                        | 14588 |
| <i>IL-PCSR: Legal Corpus for Prior Case and Statute Retrieval</i>                                                                               |       |
| Shounak Paul, Dhananjay Ghumare, Pawan Goyal, Saptarshi Ghosh and Ashutosh Modi . .                                                             | 14600 |
| <i>ESGenius: Benchmarking LLMs on Environmental, Social, and Governance (ESG) and Sustainability Knowledge</i>                                  |       |
| Chaoyue He, Xin Zhou, Yi Wu, Xinjia Yu, Yan Zhang, Lei Zhang, Di Wang, Shengfei Lyu, Hong Xu, Wang Xiaoqiao, Wei Liu and Chunyan Miao . . . . . | 14624 |
| <i>How Sememic Components Can Benefit Link Prediction for Lexico-Semantic Knowledge Graphs?</i>                                                 |       |
| Hansi Wang, Yue Wang, Qiliang Liang and Yang Liu . . . . .                                                                                      | 14666 |
| <i>WISE: Weak-Supervision-Guided Step-by-Step Explanations for Multimodal LLMs in Image Classification</i>                                      |       |
| Yiwen Jiang, Deval Mehta, Siyuan Yan, Yaling Shen, Zimu Wang and Zongyuan Ge . . . . .                                                          | 14686 |
| <i>Calibration Across Layers: Understanding Calibration Evolution in LLMs</i>                                                                   |       |
| Abhinav Joshi, Areeb Ahmad and Ashutosh Modi . . . . .                                                                                          | 14698 |
| <i>The discordance between embedded ethics and cultural inference in large language models</i>                                                  |       |
| Aida Ramezani and Yang Xu . . . . .                                                                                                             | 14727 |
| <i>SSA: Semantic Contamination of LLM-Driven Fake News Detection</i>                                                                            |       |
| Cheng Xu, Nan Yan, Shuhao Guan, Yuke Mei and Tahar Kechadi . . . . .                                                                            | 14749 |
| <i>Logits-Based Finetuning</i>                                                                                                                  |       |
| Jingyao Li, Senqiao Yang, Sitong Wu, Han Shi, Chuanyang Zheng, Hong Xu and Jiaya Jia                                                            | 14764 |
| <i>STARE at the Structure: Steering ICL Exemplar Selection with Structural Alignment</i>                                                        |       |
| Jiaqian Li, Qisheng Hu, Jing Li and Wenya Wang . . . . .                                                                                        | 14777 |
| <i>PPC-GPT: Federated Task-Specific Compression of Large Language Models via Pruning and Chain-of-Thought Distillation</i>                      |       |
| Tao Fan, Guoqiang Ma, Yuanfeng Song, Lixin Fan and Qiang Yang . . . . .                                                                         | 14795 |
| <i>Efficient Beam Search for Large Language Models Using Trie-Based Decoding</i>                                                                |       |
| Brian J Chan, Mao-xun Huang, Jui-Hung Cheng, Chao-Ting Chen and Hen-Hsen Huang . .                                                              | 14807 |
| <i>Power doesn't reside in size: A Low Parameter Hybrid Language Model (HLM) for Sentiment Analysis in Code-mixed data</i>                      |       |
| Pavan Sai Balaga, Nagasamudram Karthik, Challa Vishwanath, Raksha Sharma, Rudra Murthy and Ashish Mittal . . . . .                              | 14820 |

|                                                                                                                                                       |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Evaluating Taxonomy Free Character Role Labeling (TF-CRL) in News Stories using Large Language Models</i>                                          |       |
| David G Hobson, Derek Ruths and Andrew Piper . . . . .                                                                                                | 14829 |
| <i>MIRROR: Multimodal Cognitive Reframing Therapy for Rolling with Resistance</i>                                                                     |       |
| Subin Kim, Hoonrae Kim, Jihyun Lee, Yejin Jeon and Gary Lee . . . . .                                                                                 | 14852 |
| <i>RETAIL: Towards Real-world Travel Planning for Large Language Models</i>                                                                           |       |
| Bin Deng, Yizhe Feng, Zeming Liu, Qing Wei, Xiangrong Zhu, Shuai Chen, Yuanfang Guo and Yunhong Wang . . . . .                                        | 14882 |
| <i>Unraveling Interwoven Roles of Large Language Models in Authorship Privacy: Obfuscation, Mimicking, and Verification</i>                           |       |
| Tuc Nguyen, Yifan Hu and Thai Le . . . . .                                                                                                            | 14915 |
| <i>Reward Model Perspectives: Whose Opinions Do Reward Models Reward?</i>                                                                             |       |
| Elle . . . . .                                                                                                                                        | 14932 |
| <i>FLRC: Fine-grained Low-Rank Compressor for Efficient LLM Inference</i>                                                                             |       |
| Yu-Chen Lu, Chong-Yan Chen, Chi-Chih Chang, Yu-Fang Hu and Kai-Chiang Wu . . . . .                                                                    | 14957 |
| <i>Do You Know About My Nation? Investigating Multilingual Language Models' Cultural Literacy Through Factual Knowledge</i>                           |       |
| Eshaan Tanwar, Anwoy Chatterjee, Michael Saxon, Alon Albalak, William Yang Wang and Tanmoy Chakraborty . . . . .                                      | 14968 |
| <i>CoEvo: Coevolution of LLM and Retrieval Model for Domain-Specific Information Retrieval</i>                                                        |       |
| Ang Li, Yiquan Wu, Yinghao Hu, Lizhi Qing, Shihang Wang, Chengyuan Liu, Tao Wu, Adam Jatowt, Ming Cai, Fei Wu and Kun Kuang . . . . .                 | 14992 |
| <i>Conan-Embedding-v2: Training an LLM from Scratch for Text Embeddings</i>                                                                           |       |
| Shiyu Li, Yang Tang, Ruijie Liu, Shi-Zhe Chen and Xi Chen . . . . .                                                                                   | 15012 |
| <i>Vision-and-Language Navigation with Analogical Textual Descriptions in LLMs</i>                                                                    |       |
| Yue Zhang, Tianyi Ma, Zun Wang, Yanyuan Qiao and Parisa Kordjamshidi . . . . .                                                                        | 15029 |
| <i>MUCAR: Benchmarking Multilingual Cross-Modal Ambiguity Resolution for Multimodal Large Language Models</i>                                         |       |
| Xiaolong Wang, Zhaolu Kang, Wangyuxuan Zhai, Xinyue Lou, Yunghwei Lai, Ziyue Wang, Yawen Wang, Kaiyu Huang, Yile Wang, Peng Li and Yang Liu . . . . . | 15038 |
| <i>Mind the Gap: How BabyLMs Learn Filler-Gap Dependencies</i>                                                                                        |       |
| Chi-Yun Chang, Xueyang Huang, Humaira Nasir, Shane Storks, Olawale Akingbade and Huteng Dai . . . . .                                                 | 15061 |
| <i>Paths Not Taken: Understanding and Mending the Multilingual Factual Recall Pipeline</i>                                                            |       |
| Meng Lu, Ruochen Zhang, Carsten Eickhoff and Ellie Pavlick . . . . .                                                                                  | 15078 |
| <i>BTC-SAM: Leveraging LLMs for Generation of Bias Test Cases for Sentiment Analysis Models</i>                                                       |       |
| Zsolt T. Kardkovács, Lynda Djennane, Anna Field, Boualem Benatallah, Yacine Gaci, Fabio Casati and Walid Gaaloul . . . . .                            | 15109 |
| <i>Debate-to-Detect: Reformulating Misinformation Detection as a Real-World Debate with Large Language Models</i>                                     |       |
| Chen Han, Wenzhen Zheng and Xijin Tang . . . . .                                                                                                      | 15126 |

|                                                                                                                                                                                                                        |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Controllable Memorization in LLMs via Weight Pruning</i>                                                                                                                                                            |       |
| Chenjie Ni, Zhepeng Wang, Runxue Bao, Shangqian Gao and Yanfu Zhang . . . . .                                                                                                                                          | 15142 |
| <i>Tracing L1 Interference in English Learner Writing: A Longitudinal Corpus with Error Annotations</i>                                                                                                                |       |
| Poorvi Acharya, J. Elizabeth Liebl, Dhiman Goswami, Kai North, Marcos Zampieri and Antonios Anastasopoulos . . . . .                                                                                                   | 15158 |
| <i>DCIS: Efficient Length Extrapolation of LLMs via Divide-and-Conquer Scaling Factor Search</i>                                                                                                                       |       |
| Lei Yang, Shaoyang Xu, Jianxiang Peng, Shaolin Zhu and Deyi Xiong . . . . .                                                                                                                                            | 15180 |
| <i>Who is in the Spotlight: The Hidden Bias Undermining Multimodal Retrieval-Augmented Generation</i>                                                                                                                  |       |
| Jiayu Yao, Shenghua Liu, Yiwei Wang, Lingrui Mei, Baolong Bi, Yuyao Ge, Zhecheng Li and Xueqi Cheng . . . . .                                                                                                          | 15195 |
| <i>Let's Play Across Cultures: A Large Multilingual, Multicultural Benchmark for Assessing Language Models' Understanding of Sports</i>                                                                                |       |
| Punit Kumar Singh, Nishant Kumar, Akash Ghosh, Kunal Pasad, Khushi Soni, Manisha Jaishwal, Sriparna Saha, Syukron Abu Ishaq Alfarozi, Asres Temam Abagissa, Kitsuchart Pasupa, Haiqin Yang and Jose G Moreno . . . . . | 15206 |
| <i>Multilingual Federated Low-Rank Adaptation for Collaborative Content Anomaly Detection across Multilingual Social Media Participants</i>                                                                            |       |
| Jiaxin Li, Geng Zhao and Xiaoci Zhang . . . . .                                                                                                                                                                        | 15254 |
| <i>M3Retrieve: Benchmarking Multimodal Retrieval for Medicine</i>                                                                                                                                                      |       |
| Arkadeep Acharya, Akash Ghosh, Pradeepika Verma, Kitsuchart Pasupa, Sriparna Saha and Dr Priti Singh . . . . .                                                                                                         | 15275 |
| <i>The Hidden Strength of Disagreement: Unraveling the Consensus-Diversity Tradeoff in Adaptive Multi-Agent Systems</i>                                                                                                |       |
| Zengqing Wu and Takayuki Ito . . . . .                                                                                                                                                                                 | 15289 |
| <i>Friend or Foe? A Computational Investigation of Semantic False Friends across Romance Languages</i>                                                                                                                 |       |
| Ana Sabina Uban, Liviu P Dinu, Ioan-Bogdan Iordache, Simona Georgescu and Claudia Vlad . . . . .                                                                                                                       | 15310 |
| <i>KLAAD: Refining Attention Mechanisms to Reduce Societal Bias in Generative Language Models</i>                                                                                                                      |       |
| Seorin Kim, Dongyoung Lee and Jaejin Lee . . . . .                                                                                                                                                                     | 15325 |
| <i>SeMob: Semantic Synthesis for Dynamic Urban Mobility Prediction</i>                                                                                                                                                 |       |
| Runfei Chen, Shuyang Jiang and Wei Huang . . . . .                                                                                                                                                                     | 15347 |
| <i>DyePack: Provably Flagging Test Set Contamination in LLMs Using Backdoors</i>                                                                                                                                       |       |
| Yize Cheng, Wenxiao Wang, Mazda Moayeri and Soheil Feizi . . . . .                                                                                                                                                     | 15368 |
| <i>Minimal, Local, and Robust: Embedding-Only Edits for Implicit Bias in T2I Models</i>                                                                                                                                |       |
| Feng He, Chao Zhang and Zhixue Zhao . . . . .                                                                                                                                                                          | 15386 |
| <i>Journalism-Guided Agentic In-context Learning for News Stance Detection</i>                                                                                                                                         |       |
| Dahyun Lee, Jonghyeon Choi, Jiyoung Han and Kunwoo Park . . . . .                                                                                                                                                      | 15405 |
| <i>Less Is MuRE: Revisiting Shallow Knowledge Graph Embeddings</i>                                                                                                                                                     |       |
| Victor Charpenay and Steven Schockaert . . . . .                                                                                                                                                                       | 15429 |
| <i>Jailbreak LLMs through Internal Stance Manipulation</i>                                                                                                                                                             |       |
| Shuangjie Fu, Du Su, Beining Huang, Fei Sun, Jingang Wang, Wei Chen, Huawei Shen and Xueqi Cheng . . . . .                                                                                                             | 15456 |

|                                                                                                                                   |       |
|-----------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Pierce the Mists, Greet the Sky: Decipher Knowledge Overshadowing via Knowledge Circuit Analysis</i>                           |       |
| Haoming Huang, Yibo Yan, Jiahao Huo, Xin Zou, Xinfeng Li, Kun Wang and Xuming Hu                                                  | 15472 |
| <i>Complex Numerical Reasoning with Numerical Semantic Pre-training Framework</i>                                                 |       |
| Jun Zhang, Haihong E, Tianyi Hu, Yifan Zhu, Meina Song and Haoran Luo                                                             | 15492 |
| <i>Automated Knowledge Graph Construction using Large Language Models and Sentence Complexity Modelling</i>                       |       |
| Sydney Anuyah, Mehedi Mahmud Kaushik, Sri Rama Krishna Reddy Dwarampudi, Rakesh Shiradkar, Arjan Durresi and Sunandan Chakraborty | 15527 |
| <i>OntologyRAG-Q: Resource Development and Benchmarking for Retrieval-Augmented Question Answering in Qur'anic Tafsir</i>         |       |
| Sadam Al-Azani, Maad Alowaifeer, Alhanoof Alhunief and Ahmed Abdelali                                                             | 15552 |
| <i>The Practical Impacts of Theoretical Constructs on Empathy Modeling</i>                                                        |       |
| Allison Lahnama, Charles Welch, David Jurgens and Lucie Flek                                                                      | 15571 |
| <i>RecBase: Generative Foundation Model Pretraining for Zero-Shot Recommendation</i>                                              |       |
| Sashuai Zhou, Weinan Gan, Qijiong Liu, Ke Lei, Jieming Zhu, Hai Huang, Yan Xia, Ruiming Tang, Zhenhua Dong and Zhou Zhao          | 15599 |
| <i>Grouping Entities with Shared Properties using Multi-Facet Prompting and Property Embeddings</i>                               |       |
| Amit Gajbhiye, Thomas Bailleux, Zied Bouraoui, Luis Espinosa-Anke and Steven Schockaert                                           | 15612 |
| <i>Context-Aware Hierarchical Taxonomy Generation for Scientific Papers via LLM-Guided Multi-Aspect Clustering</i>                |       |
| Kun Zhu, Lizi Liao, Yuxuan Gu, Lei Huang, Xiaocheng Feng and Bing Qin                                                             | 15628 |
| <i>Benchmark Profiling: Mechanistic Diagnosis of LLM Benchmarks</i>                                                               |       |
| Dongjun Kim, Gyuho Shim, Yongchan Chun, Minhyuk Kim, Chanjun Park and Heuseok Lim                                                 | 15647 |
| <i>TreeReview: A Dynamic Tree of Questions Framework for Deep and Efficient LLM-based Scientific Peer Review</i>                  |       |
| Yuan Chang, Ziyue Li, Hengyuan Zhang, Yuanbo Kong, Yanru Wu, Hayden Kwok-Hay So, Zhi-jiang Guo, Liya Zhu and Ngai Wong            | 15663 |
| <i>Improving Chemical Understanding of LLMs via SMILES Parsing</i>                                                                |       |
| Yunhui Jang, Jaehyung Kim and Sungsoo Ahn                                                                                         | 15695 |
| <i>Can Large Language Models Tackle Graph Partitioning?</i>                                                                       |       |
| Yiheng Wu, Ningchao Ge, Yanmin Li, Liwei Qian, Mengna Zhu, Haoyu Yang, Haiwen Chen and Jibing Wu                                  | 15711 |
| <i>To See a World in a Spark of Neuron: Disentangling Multi-Task Interference for Training-Free Model Merging</i>                 |       |
| Zitao Fang, Guodong DU, Shuyang Yu, Yifei Guo, Yiwei Zhang, Yiyao Cao, Jing Li, Ho-Kin Tang and Sim Kuan Goh                      | 15732 |
| <i>What You Read Isn't What You Hear: Linguistic Sensitivity in Deepfake Speech Detection</i>                                     |       |
| Binh Nguyen, Shuju Shi, Ryan Ofman and Thai Le                                                                                    | 15753 |
| <i>Task-Aware Resolution Optimization for Visual Large Language Models</i>                                                        |       |
| Weiying Luo, Zhen Tan, Yifan Li, Xinyu Zhao, Kwonjoon Lee, Behzad Dariush and Tianlong Chen                                       | 15768 |

|                                                                                                                           |       |
|---------------------------------------------------------------------------------------------------------------------------|-------|
| <i>CheckEval: A reliable LLM-as-a-Judge framework for evaluating text generation using checklists</i>                     |       |
| Yukyung Lee, JoongHoon Kim, Jaehee Kim, Hyowon Cho, Jaewook Kang, Pilsung Kang and Najoung Kim .....                      | 15783 |
| <i>A Necessary Step toward Faithfulness: Measuring and Improving Consistency in Free-Text Explanations</i>                |       |
| Lingjun Zhao and Hal Daumé Iii .....                                                                                      | 15811 |
| <i>Boosting Multi-modal Keyphrase Prediction with Dynamic Chain-of-Thought in Vision-Language Models</i>                  |       |
| Qihang Ma, Shengyu Li, Jie Tang, Dingkan Yang, Chenshaodong, Yingyi Zhang, Chao Feng and Ran Jiao .....                   | 15826 |
| <i>Chart2Code53: A Large-Scale Diverse and Complex Dataset for Enhancing Chart-to-Code Generation</i>                     |       |
| Tianhao Niu, Yiming Cui, Baoxin Wang, Xiao Xu, Xin Yao, Qingfu Zhu, Dayong Wu, Shijin Wang and Wanxiang Che .....         | 15840 |
| <i>The State of Multilingual LLM Safety Research: From Measuring The Language Gap To Mitigating It</i>                    |       |
| Zheng Xin Yong, Beyza Ermis, Marzieh Fadaee, Stephen Bach and Julia Kreutzer .....                                        | 15857 |
| <i>AIP: Subverting Retrieval-Augmented Generation via Adversarial Instructional Prompt</i>                                |       |
| Saket Sanjeev Chaturvedi, Gaurav Bagwe, Lan Emily Zhang and Xiaoyong Yuan .....                                           | 15873 |
| <i>From Capabilities to Performance: Evaluating Key Functional Properties of LLM Architectures in Penetration Testing</i> |       |
| Lanxiao Huang, Daksh Dave, Tyler Cody, Peter A. Beling and Ming Jin .....                                                 | 15891 |
| <i>Editing Across Languages: A Survey of Multilingual Knowledge Editing</i>                                               |       |
| Nadir Durrani, Basel Mousi and Fahim Dalvi .....                                                                          | 15918 |
| <i>Your RAG is Unfair: Exposing Fairness Vulnerabilities in Retrieval-Augmented Generation via Back-door Attacks</i>      |       |
| Gaurav Bagwe, Saket Sanjeev Chaturvedi, Xiaolong Ma, Xiaoyong Yuan, Kuang-Ching Wang and Lan Emily Zhang .....            | 15931 |
| <i>Drift-Adapter: A Practical Approach to Near Zero-Downtime Embedding Model Upgrades in Vector Databases</i>             |       |
| Harshil Vejendla .....                                                                                                    | 15950 |
| <i>The Staircase of Ethics: Probing LLM Value Priorities through Multi-Step Induction to Complex Moral Dilemmas</i>       |       |
| Ya Wu, Qiang Sheng, Danding Wang, Guang Yang, Yifan Sun, Zhengjia Wang, Yuyan Bu and Juan Cao .....                       | 15962 |
| <i>SliceMoE: Routing Embedding Slices Instead of Tokens for Fine-Grained and Balanced Transformer Scaling</i>             |       |
| Harshil Vejendla .....                                                                                                    | 15983 |
| <i>ReSo: A Reward-driven Self-organizing LLM-based Multi-Agent System for Reasoning Tasks</i>                             |       |
| Heng Zhou, Hejia Geng, Xiangyuan Xue, Li Kang, Yiran Qin, Zhiyong Wang, Zhenfei Yin and Lei Bai .....                     | 15991 |
| <i>ConstraintLLM: A Neuro-Symbolic Framework for Industrial-Level Constraint Programming</i>                              |       |
| Weichun Shi, Minghao Liu, Wanting Zhang, Langchen Shi, Fuqi Jia, Feifei Ma and Jian Zhang                                 | 16011 |



|                                                                                                                                                                                                                                                             |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>VisEscape: A Benchmark for Evaluating Exploration-driven Decision-making in Virtual Escape Rooms</i><br>Seungwon Lim, Sungwoong Kim, Jihwan Yu, Sungjae Lee, Jiwan Chung and Youngjae Yu                                                                 | 16032 |
| <i>ESC-Judge: A Framework for Comparing Emotional Support Conversational Agents</i><br>Navid Madani and Rohini Srihari                                                                                                                                      | 16060 |
| <i>Neuron-Level Differentiation of Memorization and Generalization in Large Language Models</i><br>Ko-Wei Huang, Yi-Fu Fu, Ching-Yu Tsai, Yu-Chieh Tu, Tzu-ling Cheng, Cheng-Yu Lin, Yi-Ting Yang, Heng-Yi Liu, Keng-Te Liao, Da-Cheng Juan and Shou-De Lin | 16078 |
| <i>Sparse Neurons Carry Strong Signals of Question Ambiguity in LLMs</i><br>Zhuoxuan Zhang, Jinhao Duan, Edward Kim and Kaidi Xu                                                                                                                            | 16093 |
| <i>Do Slides Help? Multi-modal Context for Automatic Transcription of Conference Talks</i><br>Supriti Sinhamahapatra and Jan Niehues                                                                                                                        | 16112 |
| <i>Promote, Suppress, Iterate: How Language Models Answer One-to-Many Factual Queries</i><br>Tianyi Lorena Yan and Robin Jia                                                                                                                                | 16123 |
| <i>Out of Sight, Not Out of Context? Egocentric Spatial Reasoning in VLMs Across Disjoint Frames</i><br>Sahithya Ravi, Gabriel Herbert Sarch, Vibhav Vineet, Andrew D Wilson and Balasaravanan Thoravi Kumaravel                                            | 16147 |
| <i>Enhancing Chain-of-Thought Reasoning via Neuron Activation Differential Analysis</i><br>Yiru Tang, Kun Zhou, Yingqian Min, Xin Zhao, Jing Sha, Zhichao Sheng and Shijin Wang                                                                             | 16163 |
| <i>PakBBQ: A Culturally Adapted Bias Benchmark for QA</i><br>Abdullah Hashmat, Muhammad Arham Mirza and Agha Ali Raza                                                                                                                                       | 16172 |
| <i>MULTIGUARD: An Efficient Approach for AI Safety Moderation Across Languages and Modalities</i><br>Sahil Verma, Keegan Hines, Jeff Bilmes, Charlotte Siska, Luke Zettlemoyer, Hila Gonen and Chandan Singh                                                | 16185 |
| <i>Comparing human and LLM politeness strategies in free production</i><br>Haoran Zhao and Robert D. Hawkins                                                                                                                                                | 16200 |
| <i>ASTRA: A Negotiation Agent with Adaptive and Strategic Reasoning via Tool-integrated Action for Dynamic Offer Optimization</i><br>Deuksin Kwon, Jiwon Hae, Emma Clift, Daniel Shamsoddini, Jonathan Gratch and Gale Lucas                                | 16229 |
| <i>CARMA: Enhanced Compositionality in LLMs via Advanced Regularisation and Mutual Information Alignment</i><br>Nura Aljaafari, Danilo Carvalho and Andre Freitas                                                                                           | 16251 |
| <i>MEPT: Mixture of Expert Prompt Tuning as a Manifold Mapper</i><br>Runjia Zeng, Guangyan Sun, Qifan Wang, Tong Geng, Sohail Dianat, Xiaotian Han, Raghuveer Rao, Xueling Zhang, Cheng Han, Lifu Huang and Dongfang Liu                                    | 16272 |
| <i>KG-CQR: Leveraging Structured Relation Representations in Knowledge Graphs for Contextual Query Retrieval</i><br>Chi Minh Bui, Ngoc Mai Thieu, Vinh Van Nguyen, Jason J. Jung and Khac-Hoai Nam Bui                                                      | 16293 |
| <i>SABER: Uncovering Vulnerabilities in Safety Alignment via Cross-Layer Residual Connection</i><br>Maithili Joshi, Palash Nandi and Tanmoy Chakraborty                                                                                                     | 16311 |
| <i>When Truthful Representations Flip Under Deceptive Instructions?</i><br>Xianxuan Long, Yao Fu, Runchao Li, Mu Sheng, Haotian Yu, Xiaotian Han and Pan Li                                                                                                 | 16327 |

|                                                                                                                                                                                                                |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Can LLMs simulate the same correct solutions to free-response math problems as real students?</i>                                                                                                           |       |
| Yuya Asano, Diane Litman and Erin Walker . . . . .                                                                                                                                                             | 16348 |
| <i>Evaluating Behavioral Alignment in Conflict Dialogue: A Multi-Dimensional Comparison of LLM Agents and Humans</i>                                                                                           |       |
| Deuksin Kwon, Kaleen Shrestha, Bin Han, Elena Hayoung Lee and Gale Lucas . . . . .                                                                                                                             | 16378 |
| <i>RECALL: REpresentation-aligned Catastrophic-forgetting ALleviation via Hierarchical Model Merging</i>                                                                                                       |       |
| Bowen Wang, Haiyuan Wan, Liwen Shi, Chen Yang, Peng He, Yue MA, Haochen Han, Wenhao Li, Tiao Tan, Yongjian Li, Fangming Liu, Gong Yifan and Sheng Zhang . . . . .                                              | 16393 |
| <i>Not-Just-Scaling Laws: Towards a Better Understanding of the Downstream Impact of Language Model Design Decisions</i>                                                                                       |       |
| Emmy Liu, Amanda Bertsch, Lintang Sutawika, Lindia Tjauatja, Patrick Fernandes, Lara Marinov, Michael Chen, Shreya Singhal, Carolin Lawrence, Aditi Raghunathan, Kiril Gashteovski and Graham Neubig . . . . . | 16408 |
| <i>Synthetic Socratic Debates: Examining Persona Effects on Moral Decision and Persuasion Dynamics</i>                                                                                                         |       |
| Jiarui Liu, Yueqi Song, Yunze Xiao, Mingqian Zheng, Lindia Tjauatja, Jana Schaich Borg, Mona T. Diab and Maarten Sap . . . . .                                                                                 | 16440 |
| <i>Linear-Time Demonstration Selection for In-Context Learning via Gradient Estimation</i>                                                                                                                     |       |
| Ziniu Zhang, Zhenshuo Zhang, Dongyue Li, Lu Wang, Jennifer Dy and Hongyang R. Zhang                                                                                                                            | 16471 |
| <i>Speech Vecalign: an Embedding-based Method for Aligning Parallel Speech Documents</i>                                                                                                                       |       |
| Chutong Meng and Philipp Koehn . . . . .                                                                                                                                                                       | 16490 |
| <i>TurBLiMP: A Turkish Benchmark of Linguistic Minimal Pairs</i>                                                                                                                                               |       |
| Ezgi Başar, Francesca Padovani, Jaap Jumelet and Arianna Bisazza . . . . .                                                                                                                                     | 16507 |
| <i>DynamicNER: A Dynamic, Multilingual, and Fine-Grained Dataset for LLM-based Named Entity Recognition</i>                                                                                                    |       |
| Hanjun Luo, Yingbin Jin, Yiran Wang, Xinfeng Li, Tong Shang, Xuecheng Liu, Ruizhe Chen, Kun Wang, Hanan Salam, Qingsong Wen and Zuozhu Liu . . . . .                                                           | 16523 |
| <i>Reliable and Cost-Effective Exploratory Data Analysis via Graph-Guided RAG</i>                                                                                                                              |       |
| Mossad Helali, Yutai Luo, Tae Jun Ham, Jim Plotts, Ashwin Chaugule, Jichuan Chang, Parthasarathy Ranganathan and Essam Mansour . . . . .                                                                       | 16548 |
| <i>Med-PRM: Medical Reasoning Models with Stepwise, Guideline-verified Process Rewards</i>                                                                                                                     |       |
| Jaehoon Yun, Jiwoong Sohn, Jungwoo Park, Hyunjae Kim, Xiangru Tang, Daniel Shao, Yong Hoe Koo, Ko Minhyeok, Qingyu Chen, Mark Gerstein, Michael Moor and Jaewoo Kang . . . . .                                 | 16566 |
| <i>Graders Should Cheat: Privileged Information Enables Expert-Level Automated Evaluations</i>                                                                                                                 |       |
| Jin Peng Zhou, Séb Arnold, Nan Ding, Kilian Q Weinberger, Nan Hua and Fei Sha . . . . .                                                                                                                        | 16584 |
| <i>SAMULE: Self-Learning Agents Enhanced by Multi-level Reflection</i>                                                                                                                                         |       |
| Yubin Ge, Salvatore Romeo, Jason Cai, Monica Sunkara and Yi Zhang . . . . .                                                                                                                                    | 16603 |
| <i>Database-Augmented Query Representation for Information Retrieval</i>                                                                                                                                       |       |
| Soyeong Jeong, Jinheon Baek, Sukmin Cho, Sung Ju Hwang and Jong C. Park . . . . .                                                                                                                              | 16623 |
| <i>The Enemy from Within: A Study of Political Delegitimization Discourse in Israeli Political Speech</i>                                                                                                      |       |
| Naama Rivlin-Angert and Guy Mor-Lan . . . . .                                                                                                                                                                  | 16646 |



|                                                                                                                                                    |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Attention Eclipse: Manipulating Attention to Bypass LLM Safety-Alignment</i>                                                                    |       |
| Pedram Zaree, Md Abdullah Al Mamun, Quazi Mishkatul Alam, Yue Dong, Ihsen Alouani and Nael Abu-Ghazaleh .....                                      | 16660 |
| <i>Representation Potentials of Foundation Models for Multimodal Alignment: A Survey</i>                                                           |       |
| Jianglin Lu, Hailing Wang, Yi Xu, Yizhou Wang, Kuo Yang and Yun Fu .....                                                                           | 16681 |
| <i>Draft Model Knows When to Stop: Self-Verification Speculative Decoding for Long-Form Generation</i>                                             |       |
| Ziying Zhang, Jiahao Xu, Tian Liang, Xingyu Chen, Zhiwei He, Rui Wang and Zhaopeng Tu                                                              | 16697 |
| <i>Visual-Aware Speech Recognition for Noisy Scenarios</i>                                                                                         |       |
| Balaji Darur and Karan Singla .....                                                                                                                | 16710 |
| <i>Advancing Arabic Diacritization: Improved Datasets, Benchmarking, and State-of-the-Art Models</i>                                               |       |
| Abubakr Mohamed and Hamdy Mubarak .....                                                                                                            | 16719 |
| <i>Implicit Values Embedded in How Humans and LLMs Complete Subjective Everyday Tasks</i>                                                          |       |
| Arjun Arunasalam, Madison Pickering, Z. Berkay Celik and Blase Ur .....                                                                            | 16732 |
| <i>Dynamic Retriever for In-Context Knowledge Editing via Policy Optimization</i>                                                                  |       |
| Mahmud Wasif Nafee, Maiqi Jiang, Haipeng Chen and Yanfu Zhang .....                                                                                | 16756 |
| <i>LVLMS are Bad at Overhearing Human Referential Communication</i>                                                                                |       |
| Zhengxiang Wang, Weiling Li, Panagiotis Kaliosis, Owen Rambow and Susan Brennan ...                                                                | 16770 |
| <i>Let's Reason Formally: Natural-Formal Hybrid Reasoning Enhances LLM's Math Capability</i>                                                       |       |
| Ruida Wang, Yuxin Li, Yi R. Fung and Tong Zhang .....                                                                                              | 16795 |
| <i>TORSO: Template-Oriented Reasoning Towards General Tasks</i>                                                                                    |       |
| Minhyuk Kim, Seungyoon Lee and Heuiseok Lim .....                                                                                                  | 16822 |
| <i>Prototypical Human-AI Collaboration Behaviors from LLM-Assisted Writing in the Wild</i>                                                         |       |
| Sheshera Mysore, Debarati Das, Hancheng Cao and Bahareh Sarrafzadeh .....                                                                          | 16831 |
| <i>WildScore: Benchmarking MLLMs in-the-Wild Symbolic Music Reasoning</i>                                                                          |       |
| Gagan Mundada, Yash Vishe, Amit Namburi, Xin Xu, Zachary Novack, Julian McAuley and Junda Wu .....                                                 | 16859 |
| <i>TRIAL: Token Relations and Importance Aware Late-interaction for Accurate Text Retrieval</i>                                                    |       |
| Hyukkyu Kang, Injung Kim and Wook-Shin Han .....                                                                                                   | 16876 |
| <i>Do Large Language Models excel in Complex Logical Reasoning with Formal Language?</i>                                                           |       |
| Jin Jiang, Jianing Wang, Yuchen Yan, Yang Liu, Jianhua Zhu, Mengdi Zhang and Liangcai Gao                                                          | 16890 |
| <i>Fair or Framed? Political Bias in News Articles Generated by LLMs</i>                                                                           |       |
| Junho Yoo and Youhyun Shin .....                                                                                                                   | 16916 |
| <i>ReviewRL: Towards Automated Scientific Review with RL</i>                                                                                       |       |
| Sihang Zeng, Kai Tian, Kaiyan Zhang, Yuru Wang, Junqi Gao, Runze Liu, Sa Yang, Jingxuan Li, Xinwei Long, Jiaheng Ma, Bqing Qi and Bowen Zhou ..... | 16943 |
| <i>Grammar Pruning: Enabling Low-Latency Zero-Shot Task-Oriented Language Models for Edge AI</i>                                                   |       |
| Octavian Alexandru Trifan, Jason Lee Weber, Marc Titus Trifan, Alexandru Nicolau and Alexander Veidenbaum .....                                    | 16956 |
| <i>Calibrating LLMs for Text-to-SQL Parsing by Leveraging Sub-clause Frequencies</i>                                                               |       |
| Terrance Liu, Shuyi Wang, Daniel Preotiu-Pietro, Yash Chandarana and Chirag Gupta ...                                                              | 16970 |

|                                                                                                                                                                       |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>REACT: Representation Extraction And Controllable Tuning to Overcome Overfitting in LLM Knowledge Editing</i>                                                      |       |
| Haitian Zhong, Yuhuan Liu, Ziyang Xu, Guofan Liu, Qiang Liu, Shu Wu, Zhe Zhao, Liang Wang and Tieniu Tan .....                                                        | 16995 |
| <i>ThinkEdit: Interpretable Weight Editing to Mitigate Overly Short Thinking in Reasoning Models</i>                                                                  |       |
| Chung-En Sun, Ge Yan and Tsui-Wei Weng .....                                                                                                                          | 17013 |
| <i>Incorporating Diverse Perspectives in Cultural Alignment: Survey of Evaluation Benchmarks Through A Three-Dimensional Framework</i>                                |       |
| Meng-Chen Wu, Si-Chi Chin, Tess Wood, Ayush Goyal and Narayanan Sadagopan .....                                                                                       | 17038 |
| <i>Are Large Language Models Chronically Online Surfers? A Dataset for Chinese Internet Meme Explanation</i>                                                          |       |
| Yubo Xie, Chenkai Wang, Zongyang Ma and Fahui Miao .....                                                                                                              | 17074 |
| <i>RoDEval: A Robust Word Sense Disambiguation Evaluation Framework for Large Language Models</i>                                                                     |       |
| Luyang Zhang, Shuaimin Li, Yishuo Li, Kunpeng Kang, Kaiyuan Zhang, Cong Wang and Wenpeng Lu .....                                                                     | 17096 |
| <i>PychoAgent: Psychology-driven LLM Agents for Explainable Panic Prediction on Social Media during Sudden Disaster Events</i>                                        |       |
| Mengzhu Liu, Zhengqiu Zhu, Chuan Ai, Chen Gao, Xinghong Li, Lingnan He, Kaisheng Lai, Yingfeng Chen, Xin Lu, Yong Li and Quanjun Yin .....                            | 17128 |
| <i>Stepwise Reasoning Checkpoint Analysis: A Test Time Scaling Method to Enhance LLMs' Reasoning</i>                                                                  |       |
| Zezhong Wang, Xingshan Zeng, Weiwen Liu, Yufei Wang, Liangyou Li, Yasheng Wang, Lifeng Shang, Xin Jiang, Qun Liu and Kam-Fai Wong .....                               | 17147 |
| <i>Inter-sentence Context Modeling and Structure-aware Representation Enhancement for Conversational Sentiment Quadruple Extraction</i>                               |       |
| Yu Zhang, Zhaoman Zhong and Huihui LV .....                                                                                                                           | 17161 |
| <i>Igniting Creative Writing in Small Language Models: LLM-as-a-Judge versus Multi-Agent Refined Rewards</i>                                                          |       |
| Xiaolong Wei, Bo Lu, Xingyu Zhang, Zhejun Zhao, Dongdong Shen, Long Xia and Dawei Yin                                                                                 | 17172 |
| <i>Governance in Motion: Co-evolution of Constitutions and AI models for Scalable Safety</i>                                                                          |       |
| Chenhao Huang, Ziyu Shen, Yicong Ren, Huiyuan Zheng, Jiazheng Zhang, Mingxu Chai, Ming Zhang, Shihan Dou, Fan Mo, Jie Shi, Tao Gui, Qi Zhang and Xuanjing Huang ..... | 17199 |
| <i>Web Intellectual Property at Risk: Preventing Unauthorized Real-Time Retrieval by Large Language Models</i>                                                        |       |
| Yisheng Zhong, Yizhu Wen, Junfeng Guo, Mehran Kafai, Heng Huang, Hanqing Guo and Zhuangdi Zhu .....                                                                   | 17223 |
| <i>SciEvent: Benchmarking Multi-domain Scientific Event Extraction</i>                                                                                                |       |
| Bofu Dong, Pritesh Shah, Sumedh Sonawane, Tiyaasha Banerjee, Erin Brady, Xinya Du and Ming Jiang .....                                                                | 17237 |
| <i>Media Source Matters More Than Content: Unveiling Political Bias in LLM-Generated Citations</i>                                                                    |       |
| Sunhao Dai, Zhanshuo Cao, Wenjie Wang, Liang Pang, Jun Xu, See-Kiong Ng and Tat-Seng Chua                                                                             | 17268 |

|                                                                                                                                                     |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>RJE: A Retrieval-Judgment-Exploration Framework for Efficient Knowledge Graph Question Answering with LLMs</i>                                   |       |
| Can Lin, Zhengwang Jiang, Ling Zheng, Qi Zhao, Yuhang Zhang, Qi Song and Wangqiu Zhou                                                               | 17289 |
| <i>Bias Mitigation or Cultural Commonsense? Evaluating LLMs with a Japanese Dataset</i>                                                             |       |
| Taisei Yamamoto, Ryoma Kumon, Danushka Bollegala and Hitomi Yanaka . . . . .                                                                        | 17307 |
| <i>Chameleon LLMs: User Personas Influence Chatbot Personality Shifts</i>                                                                           |       |
| Jane Xing, Tianyi Niu and Shashank Srivastava . . . . .                                                                                             | 17326 |
| <i>GuessingGame: Measuring the Informativeness of Open-Ended Questions in Large Language Models</i>                                                 |       |
| Dylan Hutson, Daniel Vennemeyer, Aneesh Deshmukh, Justin Zhan and Tianyu Jiang . . .                                                                | 17345 |
| <i>SynC-LLM: Generation of Large-Scale Synthetic Circuit Code with Hierarchical Language Models</i>                                                 |       |
| Shang Liu, Yao Lu, Wenji Fang, Jing Wang and Zhiyao Xie . . . . .                                                                                   | 17362 |
| <i>Why Stop at One Error? Benchmarking LLMs as Data Science Code Debuggers for Multi-Hop and Multi-Bug Errors</i>                                   |       |
| Zhiyu Yang, Shuo Wang, Yukun Yan and Yang Deng . . . . .                                                                                            | 17378 |
| <i>Dovetail: A CPU/GPU Heterogeneous Speculative Decoding for LLM inference</i>                                                                     |       |
| Libo Zhang, Zhaoning Zhang, Xubaizhou, Rui Li, Zhiliang Tian, Songzhu Mei and Dongsheng Li . . . . .                                                | 17394 |
| <i>V-SEAM: Visual Semantic Editing and Attention Modulating for Causal Interpretability of Vision-Language Models</i>                               |       |
| Qidong Wang, Junjie Hu and Ming Jiang . . . . .                                                                                                     | 17408 |
| <i>LORAXBENCH: A Multitask, Multilingual Benchmark Suite for 20 Indonesian Languages</i>                                                            |       |
| Alham Fikri Aji and Trevor Cohn . . . . .                                                                                                           | 17433 |
| <i>MiCRo: Mixture Modeling and Context-aware Routing for Personalized Preference Learning</i>                                                       |       |
| Jingyan Shen, Jiarui Yao, Rui Yang, Yifan Sun, Feng Luo, Rui Pan, Tong Zhang and Han Zhao                                                           | 17459 |
| <i>SAFE: Schema-Driven Approximate Distance Join for Efficient Knowledge Graph Querying</i>                                                         |       |
| Sangoh Lee, Sungho Park and Wook-Shin Han . . . . .                                                                                                 | 17476 |
| <i>Structured Preference Optimization for Vision-Language Long-Horizon Task Planning</i>                                                            |       |
| Xiwen Liang, Min Lin, Weiqi Ruan, Rongtao Xu, Yuecheng Liu, Jiaqi Chen, Bingqian Lin, Yuzheng Zhuang and Xiaodan Liang . . . . .                    | 17502 |
| <i>Position: LLMs Can be Good Tutors in English Education</i>                                                                                       |       |
| Jingheng Ye, Shen Wang, Deqing Zou, Yibo Yan, Kun Wang, Hai-Tao Zheng, Ruitong Liu, Zenglin Xu, Irwin King, Philip S. Yu and Qingsong Wen . . . . . | 17528 |
| <i>CLLMate: A Multimodal Benchmark for Weather and Climate Events Forecasting</i>                                                                   |       |
| Haobo Li, Zhaowei Wang, Jiachen Wang, Yueya Wang, Alexis Kai Hon Lau and Huamin Qu                                                                  | 17548 |
| <i>Extracting and Combining Abilities For Building Multi-lingual Ability-enhanced Large Language Models</i>                                         |       |
| Zhipeng Chen, Kun Zhou, Liang Song, Xin Zhao, Bingning Wang, Weipeng Chen and Ji-Rong Wen . . . . .                                                 | 17575 |

|                                                                                                                                                    |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Evaluating the Effectiveness and Scalability of LLM-Based Data Augmentation for Retrieval</i>                                                   |       |
| Pranjal A Chitale, Bishal Santra, Yashoteja Prabhu and Amit Sharma .....                                                                           | 17593 |
| <i>Temporal Referential Consistency: Do LLMs Favor Sequences Over Absolute Time References?</i>                                                    |       |
| Ashutosh Bajpai and Tanmoy Chakraborty .....                                                                                                       | 17630 |
| <i>MemeArena: Automating Context-Aware Unbiased Evaluation of Harmfulness Understanding for Multi-modal Large Language Models</i>                  |       |
| Zixin Chen, Hongzhan Lin, Kaixin Li, Ziyang Luo, Yayue Deng and Jing Ma .....                                                                      | 17649 |
| <i>Multi-perspective Analysis of Large Language Model Domain Specialization: An Experiment in Accounting Audit Procedures Generation</i>           |       |
| Yusuke Noro .....                                                                                                                                  | 17672 |
| <i>Generator-Assistant Stepwise Rollback Framework for Large Language Model Agent</i>                                                              |       |
| Xingzuo Li, Kehai Chen, Yunfei Long, Xuefeng Bai, Yong Xu and Min Zhang .....                                                                      | 17695 |
| <i>DocAgent: An Agentic Framework for Multi-Modal Long-Context Document Understanding</i>                                                          |       |
| Li Sun, Liu He, Shuyue Jia, Yangfan He and Chenyu You .....                                                                                        | 17713 |
| <i>EasyRec: Simple yet Effective Language Models for Recommendation</i>                                                                            |       |
| Xubin Ren and Chao Huang .....                                                                                                                     | 17729 |
| <i>From Automation to Autonomy: A Survey on Large Language Models in Scientific Discovery</i>                                                      |       |
| Tianshi Zheng, Zheyue Deng, Hong Ting Tsang, Weiqi Wang, Jiaxin Bai, Zihao Wang and Yangqiu Song .....                                             | 17745 |
| <i>Mapping the Minds of LLMs: A Graph-Based Analysis of Reasoning LLMs</i>                                                                         |       |
| Zhen Xiong, Yujun Cai, Zhecheng Li and Yiwei Wang .....                                                                                            | 17763 |
| <i>ViPE: Visual Perception in Parameter Space for Efficient Video-Language Understanding</i>                                                       |       |
| Shichen Lu, Tongtian Yue, Longteng Guo, Handong Li, Xingjian He, Si Liu and Jing Liu ..                                                            | 17776 |
| <i>Alignment for Efficient Tool Calling of Large Language Models</i>                                                                               |       |
| Hongshen Xu, Zihan Wang, Zichen Zhu, Lei Pan, Xingyu Chen, Shuai Fan, Lu Chen and Kai Yu                                                           | 17788 |
| <i>ToM: Leveraging Tree-oriented MapReduce for Long-Context Reasoning in Large Language Models</i>                                                 |       |
| Jiani Guo, Zuchao Li, Jie Wu, Qianren Wang, Yun Li, Lefei Zhang, Hai Zhao and Yujiu Yang                                                           | 17805 |
| <i>BANMIME : Misogyny Detection with Metaphor Explanation on Bangla Memes</i>                                                                      |       |
| Md Ayon Mia, Akm Moshir Rahman Mazumder, Khadiza Sultana Sayma, Md Fahim, Md Tahmid Hasan Fuad, Muhammad Ibrahim Khan and Akmmahbubur Rahman ..... | 17825 |
| <i>Phi: Preference Hijacking in Multi-modal Large Language Models at Inference Time</i>                                                            |       |
| Yifan Lan, Yuanpu Cao, Weitong Zhang, Lu Lin and Jinghui Chen .....                                                                                | 17852 |
| <i>Retrieval-augmented GUI Agents with Generative Guidelines</i>                                                                                   |       |
| Ran Xu, Kaixin Ma, Wenhao Yu, Hongming Zhang, Joyce C. Ho, Carl Yang and Dong Yu                                                                   | 17878 |
| <i>COAS2W: A Chinese Older-Adults Spoken-to-Written Transformation Corpus with Context Awareness</i>                                               |       |
| Chun Kang, Zhigu Qian, Zhen Fu, Jiaojiao Fu and Yangfan Zhou .....                                                                                 | 17888 |
| <i>Answer Convergence as a Signal for Early Stopping in Reasoning</i>                                                                              |       |
| Xin Liu and Lu Wang .....                                                                                                                          | 17908 |

|                                                                                                                                 |       |
|---------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>VeriFact: Enhancing Long-Form Factuality Evaluation with Refined Fact Extraction and Reference Facts</i>                     |       |
| Xin Liu, Lechen Zhang, Sheza Munir, Yiyang Gu and Lu Wang                                                                       | 17920 |
| <i>SQUAB: Evaluating LLM robustness to Ambiguous and Unanswerable Questions in Semantic Parsing</i>                             |       |
| Simone Papicchio, Luca Cagliero and Paolo Papotti                                                                               | 17938 |
| <i>Reliable Evaluation and Benchmarks for Statement Autoformalization</i>                                                       |       |
| Auguste Poiroux, Gail Weiss, Viktor Kunčák and Antoine Bosselut                                                                 | 17959 |
| <i>VisBias: Measuring Explicit and Implicit Social Biases in Vision Language Models</i>                                         |       |
| Jen-tse Huang, Jiantong Qin, Jianping Zhang, Youliang Yuan, Wenxuan Wang and Jieyu Zhao                                         | 17982 |
| <i>Less Is More? Examining Fairness in Pruned Large Language Models for Summarising Opinions</i>                                |       |
| Nannan Huang, Haytham M. Fayek and Xiuzhen Zhang                                                                                | 18006 |
| <i>AI Sees Your Location—But With A Bias Toward The Wealthy World</i>                                                           |       |
| Jingyuan Huang, Jen-tse Huang, Ziyi Liu, Xiaoyuan Liu, Wenxuan Wang and Jieyu Zhao                                              | 18031 |
| <i>Faster In-Context Learning for LLMs via N-Gram Trie Speculative Decoding</i>                                                 |       |
| Jinglin Chen, Qiwei Li, Zuchao Li, Baoyuan Qi, Liu Guoming, Haojun Ai, Hai Zhao and Ping Wang                                   | 18052 |
| <i>From Surveys to Narratives: Rethinking Cultural Value Adaptation in LLMs</i>                                                 |       |
| Farid Adilazuarda, Chen Cecilia Liu, Iryna Gurevych and Alham Fikri Aji                                                         | 18064 |
| <i>Iterative Prompt Refinement for Safer Text-to-Image Generation</i>                                                           |       |
| Jinwoo Jeon, JunHyeok Oh, Hayeong Lee and Byung-Jun Lee                                                                         | 18092 |
| <i>Language Models as Continuous Self-Evolving Data Engineers</i>                                                               |       |
| Peidong Wang, Ming Wang, Zhiming Ma, Xiaocui Yang, Shi Feng, Daling Wang, Yifei Zhang and Kaisong Song                          | 18109 |
| <i>Unilaw-R1: A Large Language Model for Legal Reasoning with Reinforcement Learning and Iterative Inference</i>                |       |
| Hua Cai, Shuang Zhao, Liang Zhang, Xuli Shen, Qing Xu, Weilin Shen, Zihao Wen and Tianke Ban                                    | 18129 |
| <i>Exploring Response Uncertainty in MLLMs: An Empirical Evaluation under Misleading Scenarios</i>                              |       |
| Yunkai Dang, Mengxi Gao, Yibo Yan, Xin Zou, Yanggan Gu, Jungang Li, Jingyu Wang, Peijie Jiang, Aiwei Liu, Jia Liu and Xuming Hu | 18144 |
| <i>Evaluating and Aligning Human Economic Risk Preferences in LLMs</i>                                                          |       |
| Jiaxin Liu, Yixuan Tang, Yi Yang and Kar Yan Tam                                                                                | 18186 |
| <i>Ensembling Prompting Strategies for Zero-Shot Hierarchical Text Classification with Large Language Models</i>                |       |
| Mingxuan Xia, Zhijie Jiang, Haobo Wang, Junbo Zhao, Tianlei Hu and Gang Chen                                                    | 18201 |
| <i>Improbable Bigrams Expose Vulnerabilities of Incomplete Tokens in Byte-Level Tokenizers</i>                                  |       |
| Eugene Jang, Kimin Lee, Jin-Woo Chung, Keuntae Park and Seungwon Shin                                                           | 18221 |
| <i>UI-Hawk: Unleashing the Screen Stream Understanding for Mobile GUI Agents</i>                                                |       |
| Jiwen Zhang, Ya-Qi Yu, Minghui Liao, WenTao Li, Jihao Wu and Zhongyu Wei                                                        | 18229 |

|                                                                                                                                                                          |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>UniDebugger: Hierarchical Multi-Agent Framework for Unified Software Debugging</i>                                                                                    |       |
| Cheryl Lee, Chunqiu Steven Xia, Longji Yang, Jen-tse Huang, Zhou Ruixing Zhu, Lingming Zhang and Michael R. Lyu .....                                                    | 18249 |
| <i>Understanding the Thinking Process of Reasoning Models: A Perspective from Schoenfeld's Episode Theory</i>                                                            |       |
| Ming Li, Nan Zhang, Chenrui Fan, Hong Jiao, Yanbin Fu, Sydney Peters, Qingshu Xu, Robert Lissitz and Tianyi Zhou .....                                                   | 18279 |
| <i>Thread: A Logic-Based Data Organization Paradigm for How-To Question Answering with Retrieval Augmented Generation</i>                                                |       |
| Kaikai An, Fangkai Yang, Liqun Li, Junting Lu, Sitao Cheng, Shuzheng Si, Lu Wang, Pu Zhao, Lele Cao, Qingwei Lin, Saravan Rajmohan, Dongmei Zhang and Baobao Chang ..... | 18301 |
| <i>Unsupervised Word-level Quality Estimation for Machine Translation Through the Lens of Annotators (Dis)agreement</i>                                                  |       |
| Gabriele Sarti, Vilém Zouhar, Malvina Nissim and Arianna Bisazza .....                                                                                                   | 18321 |
| <i>STEER-BENCH: A Benchmark for Evaluating the Steerability of Large Language Models</i>                                                                                 |       |
| Kai Chen, Zihao He, Taiwei Shi and Kristina Lerman .....                                                                                                                 | 18339 |
| <i>Combining Constrained and Unconstrained Decoding via Boosting: BoostCD and Its Application to Information Extraction</i>                                              |       |
| Marija Sakota and Robert West .....                                                                                                                                      | 18368 |
| <i>MultiLogicNMR(er): A Benchmark and Neural-Symbolic Framework for Non-monotonic Reasoning with Multiple Extensions</i>                                                 |       |
| Yeliang Xiu and Yongmei Liu .....                                                                                                                                        | 18384 |
| <i>Beyond Demographics: Enhancing Cultural Value Survey Simulation with Multi-Stage Personality-Driven Cognitive Reasoning</i>                                           |       |
| Haijiang Liu, Qiyuan Li, Chao Gao, Yong Cao, Xiangyu Xu, Xun WU, Daniel Hershcovich and Jinguang Gu .....                                                                | 18418 |
| <i>CrystalICL: Enabling In-Context Learning for Crystal Generation</i>                                                                                                   |       |
| Ruobing Wang, Qiaoyu Tan, Yili Wang, Ying Wang and Xin Wang .....                                                                                                        | 18441 |
| <i>Towards a Unified Paradigm of Concept Editing in Large Language Models</i>                                                                                            |       |
| Zhuowen Han, Xinwei Wu, Dan Shi, Renren Jin and Deyi Xiong .....                                                                                                         | 18457 |
| <i>Step-level Verifier-guided Hybrid Test-Time Scaling for Large Language Models</i>                                                                                     |       |
| Kaiyan Chang, Yonghao Shi, Chenglong Wang, Hang Zhou, Chi Hu, Xiaoqian Liu, Yingfeng Luo, Yuan Ge, Tong Xiao and Jingbo Zhu .....                                        | 18474 |
| <i>Dynamic Expert Specialization: Towards Catastrophic Forgetting-Free Multi-Domain MoE Adaptation</i>                                                                   |       |
| Junzhuo Li, Bo Wang, Xiuze Zhou and Xuming Hu .....                                                                                                                      | 18490 |
| <i>RRInf: Efficient Influence Function Estimation via Ridge Regression for Large Language Models and Text-to-Image Diffusion Models</i>                                  |       |
| Zhuozhuo Tu, Cheng Chen and Yuxuan Du .....                                                                                                                              | 18506 |
| <i>Evaluating Spatiotemporal Consistency in Automatically Generated Sewing Instructions</i>                                                                              |       |
| Luisa Geiger, Mareike Hartmann, Michael Sullivan and Alexander Koller .....                                                                                              | 18520 |
| <i>MaZO: Masked Zeroth-Order Optimization for Multi-Task Fine-Tuning of Large Language Models</i>                                                                        |       |
| Zhen Zhang, Yifan Yang, Kai Zhen, Nathan Susanj, Athanasios Mouchtaris, Siegfried Kunzmann and Zheng Zhang .....                                                         | 18538 |



|                                                                                                                              |       |
|------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Procedural Environment Generation for Tool-Use Agents</i>                                                                 |       |
| Michael Sullivan, Mareike Hartmann and Alexander Koller .....                                                                | 18556 |
| <i>FacLens: Transferable Probe for Foreseeing Non-Factuality in Fact-Seeking Question Answering of Large Language Models</i> |       |
| Yanling Wang, Haoyang Li, Hao Zou, Jing Zhang, Xinlei He, Qi Li and Ke Xu .....                                              | 18575 |
| <i>OMS: On-the-fly, Multi-Objective, Self-Reflective Ad Keyword Generation via LLM Agent</i>                                 |       |
| Bowen Chen, Zhao Wang and Shingo Takamatsu .....                                                                             | 18595 |
| <i>Med-VRAgent: A Framework for Medical Visual Reasoning-Enhanced Agents</i>                                                 |       |
| Guangfu Guo, Xiaoqian Lu and Yue Feng .....                                                                                  | 18614 |
| <i>TrojanWave: Exploiting Prompt Learning for Stealthy Backdoor Attacks on Large Audio-Language Models</i>                   |       |
| Asif Hanif, Maha Tufail Agro, Fahad Shamshad and Karthik Nandakumar .....                                                    | 18629 |
| <i>Can LLMs be Literary Companions?: Analysing LLMs on Bengali Figures of Speech Identification</i>                          |       |
| Sourav Das and Kripabandhu Ghosh .....                                                                                       | 18646 |
| <i>Group-SAE: Efficient Training of Sparse Autoencoders for Large Language Models via Layer Groups</i>                       |       |
| Davide Ghilardi, Federico Belotti, Marco Molinari, Tao Ma and Matteo Palmonari .....                                         | 18669 |
| <i>Retrieval over Classification: Integrating Relation Semantics for Multimodal Relation Extraction</i>                      |       |
| Lei Hei, Tingjing Liao, Peiyingxin, Yiyang Qi, Jiaqi Wang, Ruiting Li and Feiliang Ren ..                                    | 18690 |
| <i>PunMemeCN: A Benchmark to Explore Vision-Language Models' Understanding of Chinese Pun Memes</i>                          |       |
| Zhijun Xu, Siyu Yuan, Yiqiao Zhang, Jingyu Sun, Tong Zheng and Deqing Yang .....                                             | 18706 |
| <i>UltraIF: Advancing Instruction Following from the Wild</i>                                                                |       |
| Kaikai An, Li Sheng, Ganqu Cui, Shuzheng Si, Ning Ding, Yu Cheng and Baobao Chang ..                                         | 18723 |
| <i>Identifying Pre-training Data in LLMs: A Neuron Activation-Based Detection Framework</i>                                  |       |
| Hongyi Tang, Zhihao Zhu and Yi Yang .....                                                                                    | 18739 |
| <i>TreeRare: Syntax Tree-Guided Retrieval and Reasoning for Knowledge-Intensive Question Answering</i>                       |       |
| Boyi Zhang, Zhuo Liu and Hangfeng He .....                                                                                   | 18753 |
| <i>Mapping Toxic Comments Across Demographics: A Dataset from German Public Broadcasting</i>                                 |       |
| Jan Fillies, Michael Peter Hoffmann, Rebecca Reichel, Roman Salzwedel, Sven Bodemer and Adrian Paschke .....                 | 18775 |
| <i>Small Models, Big Results: Achieving Superior Intent Extraction through Decomposition</i>                                 |       |
| Danielle Cohen, Yoni Halpern, Noam Kahlon, Joel Oren, Omri Berkovitch, Sapir Caduri, Ido Dagan and Anatoly Efros .....       | 18792 |
| <i>On Pruning State-Space LLMs</i>                                                                                           |       |
| Tamer Ghattas, Michael Hassid and Roy Schwartz .....                                                                         | 18812 |
| <i>An Orthogonal High-Rank Adaptation for Large Language Models</i>                                                          |       |
| Xin Zhang, Guang-Ze Chen, Shuzhen Li, Zhulin Liu, C.L.Philip Chen and Tong Zhang ...                                         | 18827 |
| <i>BSFA: Leveraging the Subspace Dichotomy to Accelerate Neural Network Training</i>                                         |       |
| WenJie Zhou, Bohan Wang, Wei Chen and Xueqi Cheng .....                                                                      | 18846 |
| <i>Debatable Intelligence: Benchmarking LLM Judges via Debate Speech Evaluation</i>                                          |       |
| Noy Sternlicht, Ariel Gera, Roy Bar-Haim, Tom Hope and Noam Slonim .....                                                     | 18862 |

|                                                                                                                                                                                                        |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>METok: Multi-Stage Event-based Token Compression for Efficient Long Video Understanding</i>                                                                                                         |       |
| Mengyue Wang, Shuo Chen, Kristian Kersting, Volker Tresp and Yunpu Ma . . . . .                                                                                                                        | 18882 |
| <i>VisiPruner: Decoding Discontinuous Cross-Modal Dynamics for Efficient Multimodal LLMs</i>                                                                                                           |       |
| Yingqi Fan, Anhao Zhao, Jinlan Fu, Junlong Tong, Hui Su, Yijie Pan, Wei Zhang and Xiaoyu Shen . . . . .                                                                                                | 18897 |
| <i>Beyond Static Testbeds: An Interaction-Centric Agent Simulation Platform for Dynamic Recommender Systems</i>                                                                                        |       |
| Song Jin, Juntian Zhang, Yuhan Liu, Xun Zhang, Yufei Zhang, Guojun Yin, Fei Jiang, Wei Lin and Rui Yan . . . . .                                                                                       | 18915 |
| <i>SheetDesigner: MLLM-Powered Spreadsheet Layout Generation with Rule-Based and Vision-Based Reflection</i>                                                                                           |       |
| Qin Chen, Yuanyi Ren, Xiaojun Ma, Mugeng Liu, Shi Han and Dongmei Zhang . . . . .                                                                                                                      | 18933 |
| <i>CAIR: Counterfactual-based Agent Influence Ranker for Agentic AI Workflows</i>                                                                                                                      |       |
| Amit Giloni, Chiara Picardi, Roy Betser, Shamik Bose, Aishvariya Priya Rathina Sabapathy and Roman Vainshtein . . . . .                                                                                | 18952 |
| <i>ReSURE: Regularizing Supervision Unreliability for Multi-turn Dialogue Fine-tuning</i>                                                                                                              |       |
| Yiming Du, Yifan Xiang, Bin Liang, Dahua Lin, Kam-Fai Wong and Fei Tan . . . . .                                                                                                                       | 18979 |
| <i>Precise In-Parameter Concept Erasure in Large Language Models</i>                                                                                                                                   |       |
| Yoav Gur-Arieh, Clara Haya Suslik, Yihuai Hong, Fazl Barez and Mor Geva . . . . .                                                                                                                      | 18998 |
| <i>PhonoThink: Improving Large Language Models' Reasoning on Chinese Phonological Ambiguities</i>                                                                                                      |       |
| Jianfei Ma, Zhaoxin Feng, Emmanuele Chersoni, Huacheng Song and Ziqi Zhang . . . . .                                                                                                                   | 19019 |
| <i>SAFE-SQL: Self-Augmented In-Context Learning with Fine-grained Example Selection for Text-to-SQL</i>                                                                                                |       |
| Jimin Lee, Ingeol Baek, Byeongjeong Kim, Hyunkyung Bae and Hwanhee Lee . . . . .                                                                                                                       | 19035 |
| <i>ExpandR: Teaching Dense Retrievers Beyond Queries with LLM Guidance</i>                                                                                                                             |       |
| Sijia Yao, Pengcheng Huang, Zhenghao Liu, Yu Gu, Yukun Yan, Shi Yu and Ge Yu . . . . .                                                                                                                 | 19048 |
| <i>Anecdotoxing: Automated Red-Teaming Across Language and Place</i>                                                                                                                                   |       |
| Alejandro Cuevas, Saloni Dash, Bharat Kumar Nayak, Dan Vann and Madeleine I. G. Daepf . . . . .                                                                                                        | 19067 |
| <i>ACING: Actor-Critic for Instruction Learning in Black-Box LLMs</i>                                                                                                                                  |       |
| Salma Kharrat, Fares Fourati and Marco Canini . . . . .                                                                                                                                                | 19087 |
| <i>Women, Infamous, and Exotic Beings: A Comparative Study of Honorific Usages in Wikipedia and LLMs for Bengali and Hindi</i>                                                                         |       |
| Sourabrata Mukherjee, Atharva Mehta, Sougata Saha, Akhil Arora and Monojit Choudhury . . . . .                                                                                                         | 19115 |
| <i>Process-Supervised Reward Models for Verifying Clinical Note Generation: A Scalable Approach Guided by Domain Expertise</i>                                                                         |       |
| Hanyin Wang, Chufan Gao, Qiping Xu, Bolun Liu, Guleid Hussein, Hariprasad Reddy Korsapati, Mohamad El Labban, Kingsley Iheasirim, Mohamed Hassan, Gokhan Anil, Brian Bartlett and Jimeng Sun . . . . . | 19139 |
| <i>GCML: Gradient Coherence Guided Meta-Learning for Cross-Domain Emerging Topic Rumor Detection</i>                                                                                                   |       |
| Zejiang He, Jingyuan Huang, Menglong Lu, Zhen Huang, Shanshan Liu, Zhiliang Tian and Dongsheng Li . . . . .                                                                                            | 19160 |



|                                                                                                                                                          |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Can LLMs Generate and Solve Linguistic Olympiad Puzzles?</i>                                                                                          |       |
| Neh Majmudar and Elena Filatova . . . . .                                                                                                                | 19175 |
| <i>E2LLM: Encoder Elongated Large Language Models for Long-Context Understanding and Reasoning</i>                                                       |       |
| Zihan Liao, Jun Wang, Hang Yu, Lingxiao Wei, Jianguo Li, Jun Wang and Wei Zhang . . . .                                                                  | 19213 |
| <i>DivScore: Zero-Shot Detection of LLM-Generated Text in Specialized Domains</i>                                                                        |       |
| Zhihui Chen, Kai He, Yucheng Huang, Yunxiao Zhu and Mengling Feng . . . . .                                                                              | 19243 |
| <i>Multi-Document Event Extraction Using Large and Small Language Models</i>                                                                             |       |
| Qingkai Min, Zitian Qu, Qipeng Guo, Xiangkun Hu, Zheng Zhang and Yue Zhang . . . . .                                                                     | 19266 |
| <i>MA-GTS: A Multi-Agent Framework for Solving Complex Graph Problems in Real-World Applications</i>                                                     |       |
| Zike Yuan, Ming Liu, Hui Wang and Bing Qin . . . . .                                                                                                     | 19298 |
| <i>Enhancing Speech Large Language Models with Prompt-Aware Mixture of Audio Encoders</i>                                                                |       |
| Weiqiao Shan, Yuang Li, Yuhao Zhang, Yingfeng Luo, Chen Xu, Xiaofeng Zhao, Long Meng, Yunfei Lu, Min Zhang, Hao Yang, Tong Xiao and JingBo Zhu . . . . . | 19317 |
| <i>CIKT: A Collaborative and Iterative Knowledge Tracing Framework with Large Language Models</i>                                                        |       |
| Runze Li, Siyu wu, Jun Wang and Wei Zhang . . . . .                                                                                                      | 19333 |
| <i>Mitigating Hallucinations in LM-Based TTS Models via Distribution Alignment Using GFlowNets</i>                                                       |       |
| Chenlin Liu, Minghui Fang, Patrick Zhang, Wei Zhou, Jie Gao and Jiqing Han . . . . .                                                                     | 19347 |
| <i>MolErr2Fix: Benchmarking LLM Trustworthiness in Chemistry via Modular Error Detection, Localization, Explanation, and Correction</i>                  |       |
| Yuyang Wu, Jinhui Ye, Shuhao Zhang, Lu Dai, Yonatan Bisk and Olexandr Isayev . . . . .                                                                   | 19366 |
| <i>Shared Path: Unraveling Memorization in Multilingual LLMs through Language Similarities</i>                                                           |       |
| Xiaoyu Luo, Yiyi Chen, Johannes Bjerva and Qiongxiu Li . . . . .                                                                                         | 19384 |
| <i>Embedding Domain Knowledge for Large Language Models via Reinforcement Learning from Augmented Generation</i>                                         |       |
| Chaojun Nie, Jun Zhou, Guanxiang Wang, Shisong Wu and Zichen Wang . . . . .                                                                              | 19401 |
| <i>LLM-Driven Completeness and Consistency Evaluation for Cultural Heritage Data Augmentation in Cross-Modal Retrieval</i>                               |       |
| Jian Zhang, Junyi Guo, Junyi Yuan, Huanda Lu, Yanlin Zhou, Fangyu Wu, Qiufeng Wang and Dongming Lu . . . . .                                             | 19419 |
| <i>Artificial Impressions: Evaluating Large Language Model Behavior Through the Lens of Trait Impressions</i>                                            |       |
| Nicholas Deas and Kathleen McKeown . . . . .                                                                                                             | 19430 |
| <i>Mitigating Hallucinations in Large Vision-Language Models via Entity-Centric Multimodal Preference Optimization</i>                                   |       |
| Jiulong Wu, Zhengliang Shi, Shuaiqiang Wang, Jizhou Huang, Dawei Yin, Lingyong Yan, Min Cao and Min Zhang . . . . .                                      | 19457 |
| <i>3DS: Medical Domain Adaptation of LLMs via Decomposed Difficulty-based Data Selection</i>                                                             |       |
| Hongxin Ding, Yue Fang, Runchuan Zhu, Xinke Jiang, Jinyang Zhang, Yongxin Xu, Weibin Liao, Xu Chu, Junfeng Zhao and Yasha Wang . . . . .                 | 19474 |
| <i>InfiniBench: A Benchmark for Large Multi-Modal Models in Long-Form Movies and TV Shows</i>                                                            |       |
| Kirolos Ataallah, Eslam Mohamed Bakr, Mahmoud Ahmed, Chenhui Gou, Khushbu Pahwa, Jian Ding and Mohamed Elhoseiny . . . . .                               | 19497 |

|                                                                                                                                                              |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Intrinsic Test of Unlearning Using Parametric Knowledge Traces</i>                                                                                        |       |
| Yihuai Hong, Lei Yu, Haiqin Yang, Shauli Ravfogel and Mor Geva .....                                                                                         | 19525 |
| <i>Speculative Streaming: Efficient and Scalable Speculative Decoding with Multi-Stream Attention</i>                                                        |       |
| Nikhil Bhendawade, Irina Belousova, Qichen Fu, Henry Mason, Antonie Lin, Mohammad Rastegari and Mahyar Najibi .....                                          | 19548 |
| <i>Evaluating Cognitive-Behavioral Fixation via Multimodal User Viewing Patterns on Social Media</i>                                                         |       |
| Yujie Wang, Yunwei Zhao, Jing Yang, Han Han, Shiguang Shan and Jie Zhang .....                                                                               | 19572 |
| <i>Mind the Gap: A Closer Look at Tokenization for Multiple-Choice Question Answering with LLMs</i>                                                          |       |
| Mario Sanz-Guerrero, Minh Duc Bui and Katharina von der Wense .....                                                                                          | 19585 |
| <i>VocalNet: Speech LLMs with Multi-Token Prediction for Faster and High-Quality Generation</i>                                                              |       |
| Yuhao Wang, Heyang Liu, Ziyang Cheng, Ronghua Wu, Qunshan Gu, Yanfeng Wang and Yu Wang .....                                                                 | 19596 |
| <i>Path Drift in Large Reasoning Models: How First-Person Commitments Override Safety</i>                                                                    |       |
| Yuyi Huang, Runzhe Zhan, Lidia S. Chao, Ailin Tao and Derek F. Wong .....                                                                                    | 19614 |
| <i>CBP-Tuning: Efficient Local Customization for Black-box Large Language Models</i>                                                                         |       |
| Jiaxuan Zhao, Naibin Gu, Yuchen Feng, Xiyu Liu, Peng Fu, Zheng Lin and Weiping Wang .....                                                                    | 19629 |
| <i>Beyond the Score: Uncertainty-Calibrated LLMs for Automated Essay Assessment</i>                                                                          |       |
| Ahmed Karim, Qiao Wang and Zheng Yuan .....                                                                                                                  | 19643 |
| <i>Humans Hallucinate Too: Language Models Identify and Correct Subjective Annotation Errors With Label-in-a-Haystack Prompts</i>                            |       |
| Georgios Chochlakis, Peter Wu, Tikka Arjun Singh Bedi, Marcus Ma, Kristina Lerman and Shrikanth Narayanan .....                                              | 19649 |
| <i>Do It Yourself (DIY): Modifying Images for Poems in a Zero-Shot Setting Using Weighted Prompt Manipulation</i>                                            |       |
| Sofia Jamil, Kotla Sai Charan, Sriparna Saha, Koustava Goswami and Joseph K J .....                                                                          | 19669 |
| <i>Looking Beyond Text: Reducing Language Bias in Large Vision-Language Models via Multimodal Dual-Attention and Soft-Image Guidance</i>                     |       |
| Haozhe Zhao, Shuzheng Si, Liang Chen, Yichi Zhang, Maosong Sun, Baobao Chang and Minjia Zhang .....                                                          | 19678 |
| <i>Who Holds the Pen? Caricature and Perspective in LLM Retellings of History</i>                                                                            |       |
| Lubna Zahan Lamia, Mabsur Fatin Bin Hossain and Md Mosaddek Khan .....                                                                                       | 19703 |
| <i>DSMoE: Matrix-Partitioned Experts with Dynamic Routing for Computation-Efficient Dense LLMs</i>                                                           |       |
| Minxuan Lv, Zhenpeng Su, Leiyu Pan, Yizhe Xiong, Zijia Lin, Hui Chen, Wei Zhou, Jungong Han, Guiguang Ding, Wenwu Ou, Di Zhang, Kun Gai and Songlin Hu ..... | 19723 |
| <i>Search Wisely: Mitigating Sub-optimal Agentic Searches By Reducing Uncertainty</i>                                                                        |       |
| Peilin Wu, Mian Zhang, Xinlu Zhang, Xinya Du and Zhiyu Chen .....                                                                                            | 19735 |
| <i>Child-Directed Language Does Not Consistently Boost Syntax Learning in Language Models</i>                                                                |       |
| Francesca Padovani, Jaap Jumelet, Yevgen Matushevych and Arianna Bisazza .....                                                                               | 19747 |
| <i>Benchmarking Debiasing Methods for LLM-based Parameter Estimates</i>                                                                                      |       |
| Nicolas Audinet de Pieuchon, Adel Daoud, Connor Thomas Jerzak, Moa Johansson and Richard Johansson .....                                                     | 19769 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>(Almost) Free Modality Stitching of Foundation Models</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                     |       |
| Jaisidh Singh, Diganta Misra, Boris Knyazev and Antonio Orvieto . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                        | 19785 |
| <i>VERITAS: Leveraging Vision Priors and Expert Fusion to Improve Multimodal Data</i>                                                                                                                                                                                                                                                                                                                                                                                                                            |       |
| Tingqiao Xu, Ziru Zeng and Jiayu Chen . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 19802 |
| <i>Rescorla-Wagner Steering of LLMs for Undesired Behaviors over Disproportionate Inappropriate Context</i>                                                                                                                                                                                                                                                                                                                                                                                                      |       |
| Rushi Wang, Jiateng Liu, Cheng Qian, Yifan Shen, Yanzhou Pan, Zhaozhuo Xu, Ahmed Abbasi, Heng Ji and Denghui Zhang . . . . .                                                                                                                                                                                                                                                                                                                                                                                     | 19822 |
| <i>Exploring Artificial Image Generation for Stance Detection</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                |       |
| Zheng kang Zhang, Zhongqing Wang and Guodong Zhou . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 19858 |
| <i>Hope vs. Hate: Understanding User Interactions with LGBTQ+ News Content in Mainstream US News Media through the Lens of Hope Speech</i>                                                                                                                                                                                                                                                                                                                                                                       |       |
| Jonathan Pofcher, Christopher M Homan, Randall Sell and Ashiqur R. KhudaBukhsh . . . .                                                                                                                                                                                                                                                                                                                                                                                                                           | 19874 |
| <i>Flaw or Artifact? Rethinking Prompt Sensitivity in Evaluating LLMs</i>                                                                                                                                                                                                                                                                                                                                                                                                                                        |       |
| Andong Hua, Kenan Tang, Chenhe Gu, Jindong Gu, Eric Wong and Yao Qin . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                   | 19901 |
| <i>Topic Coverage-based Demonstration Retrieval for In-Context Learning</i>                                                                                                                                                                                                                                                                                                                                                                                                                                      |       |
| Wonbin Kweon, SeongKu Kang, Runchu Tian, Pengcheng Jiang, Jiawei Han and Hwanjo Yu                                                                                                                                                                                                                                                                                                                                                                                                                               | 19912 |
| <i>On the Same Wavelength? Evaluating Pragmatic Reasoning in Language Models across Broad Concepts</i>                                                                                                                                                                                                                                                                                                                                                                                                           |       |
| Linlu Qiu, Cedegao E. Zhang, Joshua B. Tenenbaum, Yoon Kim and Roger P. Levy . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                           | 19925 |
| <i>MuseScorer: Idea Originality Scoring At Scale</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                             |       |
| Ali Sarosh Bangash, Krish Veera, Ishfat Abrar Islam and Raiyan Abdul Baten . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                             | 19948 |
| <i>SAFENUDGE: Safeguarding Large Language Models in Real-time with Tunable Safety-Performance Trade-offs</i>                                                                                                                                                                                                                                                                                                                                                                                                     |       |
| Joao Fonseca, Andrew Bell and Julia Stoyanovich . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 19967 |
| <i>RaDeR: Reasoning-aware Dense Retrieval Models</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                             |       |
| Debrup Das, Sam O’Nuallain and Razieh Rahimi . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 19982 |
| <i>A Culturally-diverse Multilingual Multimodal Video Benchmark &amp; Model</i>                                                                                                                                                                                                                                                                                                                                                                                                                                  |       |
| Bhuiyan Sanjid Shafique, Ashmal Vayani, Muhammad Maaz, Hanoona Abdul Rasheed, Dinura Dissanayake, Mohammed Irfan Kurpath, Yahya Hmaiti, Go Inoue, Jean Lahoud, Md. Safirur Rashid, Shadid Intisar Quasem, Maheen Fatima, Franco Vidal, Mykola Maslych, Ketan Pravin More, Sanoojan Baliah, Hasindri Watawana, Yuhao Li, Fabian Farestam, Leon Schaller, Roman Tyntsiv, Simon Weber, Hisham Cholakkal, Ivan Laptev, Shin’ichi Satoh, Michael Felsberg, Mubarak Shah, Salman Khan and Fahad Shahbaz Khan . . . . . | 20010 |
| <i>DRES: Fake news detection by dynamic representation and ensemble selection</i>                                                                                                                                                                                                                                                                                                                                                                                                                                |       |
| Faramarz Farhangian, Leandro Augusto Ensina, George D C Cavalcanti and Rafael M. O. Cruz                                                                                                                                                                                                                                                                                                                                                                                                                         | 20035 |
| <i>A Graph-Theoretical Framework for Analyzing the Behavior of Causal Language Models</i>                                                                                                                                                                                                                                                                                                                                                                                                                        |       |
| Rashin Rahnamoun and Mehrnoush Shamsfard . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 20054 |
| <i>Membership and Memorization in LLM Knowledge Distillation</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                 |       |
| Ziqi Zhang, Ali Shahin Shamsabadi, Hanxiao Lu, Yifeng Cai and Hamed Haddadi . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                            | 20086 |

|                                                                                                                                                                     |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Balanced Multi-Factor In-Context Learning for Multilingual Large Language Models</i>                                                                             |       |
| Masahiro Kaneko, Alham Fikri Aji and Timothy Baldwin                                                                                                                | 20097 |
| <i>Efficient Context Selection for Long-Context QA: No Tuning, No Iteration, Just Adaptive-k</i>                                                                    |       |
| Chihiro Taguchi, Seiji Maekawa and Nikita Bhutani                                                                                                                   | 20117 |
| <i>Languages Still Left Behind: Toward a Better Multilingual Machine Translation Benchmark</i>                                                                      |       |
| Chihiro Taguchi, Seng Mai, Keita Kurabe, Yusuke Sakai, Georgina Agyei, Soudabeh Eslami and David Chiang                                                             | 20143 |
| <i>Think Globally, Group Locally: Evaluating LLMs Using Multi-Lingual Word Grouping Games</i>                                                                       |       |
| César Guerra-Solano, Zhuochun Li and Xiang Lorraine Li                                                                                                              | 20156 |
| <i>Pointing to a Llama and Call it a Camel: On the Sycophancy of Multimodal Large Language Models</i>                                                               |       |
| Renjie Pi, Kehao Miao, LI Peihang, Runtao Liu, Jiahui Gao, Jipeng Zhang and Xiaofang Zhou                                                                           | 20178 |
| <i>MR. Judge: Multimodal Reasoner as a Judge</i>                                                                                                                    |       |
| Renjie Pi, Haoping Bai, Qibin Chen, Xiaoming Simon Wang, Jiulong Shan, Xiaojiang Liu and Meng Cao                                                                   | 20193 |
| <i>MobiZO: Enabling Efficient LLM Fine-Tuning at the Edge via Inference Engines</i>                                                                                 |       |
| Lei Gao, Amir Ziashahabi, Yue Niu, Salman Avestimehr and Murali Annavaram                                                                                           | 20218 |
| <i>Fann or Flop: A Multigenre, Multitier Benchmark for Arabic Poetry Understanding in LLMs</i>                                                                      |       |
| Wafa Al Ghallabi, Ritesh Thawkar, Sara Ghaboura, Ketan Pravin More, Omkar Thawakar, Hisham Cholakal, Salman Khan and Rao Muhammad Anwer                             | 20236 |
| <i>CoMAT: Chain of Mathematically Annotated Thought Improves Mathematical Reasoning</i>                                                                             |       |
| Joshua Ong Jun Leang, Aryo Pradipta Gema and Shay B Cohen                                                                                                           | 20257 |
| <i>s1: Simple test-time scaling</i>                                                                                                                                 |       |
| Niklas Muennighoff, Zitong Yang, Weijia Shi, Xiang Lisa Li, Li Fei-Fei, Hannaneh Hajishirzi, Luke Zettlemoyer, Percy Liang, Emmanuel Candes and Tatsunori Hashimoto | 20287 |
| <i>Learning Subjective Label Distributions via Sociocultural Descriptors</i>                                                                                        |       |
| Mohammed Fayiz Parappan and Ricardo Henao                                                                                                                           | 20334 |
| <i>COM-BOM: Bayesian Exemplar Search for Efficiently Exploring the Accuracy-Calibration Pareto Frontier</i>                                                         |       |
| Gaoxiang Luo and Aryan Deshwal                                                                                                                                      | 20351 |
| <i>ML-Promise: A Multilingual Dataset for Corporate Promise Verification</i>                                                                                        |       |
| Yohei Seki, Hakusen Shu, Anaïs Lhuissier, Hanwool Lee, Juyeon Kang, Min-Yuh Day and Chung-Chi Chen                                                                  | 20365 |
| <i>Reading Between the Prompts: How Stereotypes Shape LLM's Implicit Personalization</i>                                                                            |       |
| Vera Neplenbroek, Arianna Bisazza and Raquel Fernández                                                                                                              | 20379 |
| <i>Paired by the Teacher: Turning Unpaired Data into High-Fidelity Pairs for Low-Resource Text Generation</i>                                                       |       |
| Yen-Ju Lu, Thomas Thebaud, Laureano Moro-Velazquez, Najim Dehak and Jesus Villalba                                                                                  | 20413 |
| <i>Please Translate Again: Two Simple Experiments on Whether Human-Like Reasoning Helps Translation</i>                                                             |       |
| Di Wu, Seth Aycok and Christof Monz                                                                                                                                 | 20436 |

|                                                                                                                                                                                                          |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>How Do Large Vision-Language Models See Text in Image? Unveiling the Distinctive Role of OCR Heads</i>                                                                                                |       |
| Ingeol Baek, Hwan Chang, Sunghyun Ryu and Hwanhee Lee . . . . .                                                                                                                                          | 20453 |
| <i>Explainability and Interpretability of Multilingual Large Language Models: A Survey</i>                                                                                                               |       |
| Lucas Resck, Isabelle Augenstein and Anna Korhonen . . . . .                                                                                                                                             | 20466 |
| <i>Decoding the Rule Book: Extracting Hidden Moderation Criteria from Reddit Communities</i>                                                                                                             |       |
| Youngwoo Kim, Himanshu Beniwal, Steven L. Johnson and Thomas Hartvigsen . . . . .                                                                                                                        | 20499 |
| <i>AcT2I: Evaluating and Improving Action Depiction in Text-to-Image Models</i>                                                                                                                          |       |
| Vatsal Malaviya, Agneet Chatterjee, Maitreya Patel, Yezhou Yang and Chitta Baral . . . . .                                                                                                               | 20511 |
| <i>Assessing French Readability for Adults with Low Literacy: A Global and Local Perspective</i>                                                                                                         |       |
| Wafa Aissa, Thibault Bañeras-Roux, Elodie Vanzeveren, Lingyun Gao, Rodrigo Wilkens and Thomas François . . . . .                                                                                         | 20529 |
| <i>LILaC: Late Interacting in Layered Component Graph for Open-domain Multimodal Multihop Retrieval</i>                                                                                                  |       |
| Joohyung Yun, Doyup Lee and Wook-Shin Han . . . . .                                                                                                                                                      | 20552 |
| <i>DiCoRe: Enhancing Zero-shot Event Detection via Divergent-Convergent LLM Reasoning</i>                                                                                                                |       |
| Tanmay Parekh, Kartik Mehta, Ninareh Mehrabi, Kai-Wei Chang and Nanyun Peng . . . . .                                                                                                                    | 20572 |
| <i>SNaRe: Domain-aware Data Generation for Low-Resource Event Detection</i>                                                                                                                              |       |
| Tanmay Parekh, Yuxuan Dong, Lucas Bandarkar, Artin Kim, I-Hung Hsu, Kai-Wei Chang and Nanyun Peng . . . . .                                                                                              | 20595 |
| <i>Table-R1: Inference-Time Scaling for Table Reasoning Tasks</i>                                                                                                                                        |       |
| Zheyuan Yang, Lyuhao Chen, Arman Cohan and Yilun Zhao . . . . .                                                                                                                                          | 20617 |
| <i>LimRank: Less is More for Reasoning-Intensive Information Reranking</i>                                                                                                                               |       |
| Tingyu Song, Yilun Zhao, Siyue Zhang, Chen Zhao and Arman Cohan . . . . .                                                                                                                                | 20637 |
| <i>PlanGEN: A Multi-Agent Framework for Generating Planning and Reasoning Trajectories for Complex Problem Solving</i>                                                                                   |       |
| Mihir Parmar, Xin Liu, Palash Goyal, Yanfei Chen, Long Le, Swaroop Mishra, Hossein Mobahi, Jindong Gu, Zifeng Wang, Hootan Nakhost, Chitta Baral, Chen-Yu Lee, Tomas Pfister and Hamid Palangi . . . . . | 20652 |
| <i>An Empirical Study on Strong-Weak Model Collaboration for Repo-level Code Generation</i>                                                                                                              |       |
| Shubham Gandhi, Atharva Naik, Yiqing Xie and Carolyn Rose . . . . .                                                                                                                                      | 20679 |
| <i>What are Foundation Models Cooking in the Post-Soviet World?</i>                                                                                                                                      |       |
| Anton Lavrouk, Tarek Naous, Alan Ritter and Wei Xu . . . . .                                                                                                                                             | 20699 |
| <i>LogiDynamics: Unraveling the Dynamics of Inductive, Abductive and Deductive Logical Inferences in LLM Reasoning</i>                                                                                   |       |
| Tianshi Zheng, Cheng Jiayang, Chunyang Li, Haochen Shi, Zihao Wang, Jiaxin Bai, Yangqiu Song, Ginny Wong and Simon See . . . . .                                                                         | 20722 |
| <i>EcoLoRA: Communication-Efficient Federated Fine-Tuning of Large Language Models</i>                                                                                                                   |       |
| Han Liu, Ruoyao Wen, Srijith Nair, Jia Liu, Wenjing Lou, Chongjie Zhang, William Yeoh, Yevgeniy Vorobeychik and Ning Zhang . . . . .                                                                     | 20744 |
| <i>Memorization <math>\neq</math> Understanding: Do Large Language Models Have the Ability of Scenario Cognition?</i>                                                                                    |       |
| Boxiang Ma, Ru Li, Wang Yuanlong, Hongye Tan and Xiaoli Li . . . . .                                                                                                                                     | 20759 |

|                                                                                                                                                         |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Priority on High-Quality: Selecting Instruction Data via Consistency Verification of Noise Injection</i>                                             |       |
| Hong Zhang, Feng Zhao, Ruilin Zhao, Cheng Yan and Kangzheng Liu .....                                                                                   | 20776 |
| <i>Can Prompts Rewind Time for LLMs? Evaluating the Effectiveness of Prompted Knowledge Cutoffs</i>                                                     |       |
| Xin Gao, Ruiyi Zhang, Daniel Du, Saurabh Mahindre, Sai Ashish Somayajula and Pengtao Xie                                                                | 20789 |
| <i>DSVD: Dynamic Self-Verify Decoding for Faithful Generation in Large Language Models</i>                                                              |       |
| YiQiu Guo, Yuchen Yang, Zhe Chen, Pingjie Wang, Yusheng Liao, Ya Zhang, Yanfeng Wang and Yu Wang .....                                                  | 20801 |
| <i>Metric Calculating Benchmark: Code-Verifiable Complicate Instruction Following Benchmark for Large Language Models</i>                               |       |
| Hyeonseok Moon, Seongtae Hong, Jaehyung Seo and Heuseok Lim .....                                                                                       | 20821 |
| <i>Generative Annotation for ASR Named Entity Correction</i>                                                                                            |       |
| Yuanchang Luo, Daimeng Wei, Shaojun Li, Hengchao Shang, Jiaxin Guo, Zongyao Li, Zhanglin Wu, Xiaoyu Chen, Zhiqiang Rao, Jinlong Yang and Hao Yang ..... | 20836 |
| <i>SOLAR: Towards Characterizing Subjectivity of Individuals through Modeling Value Conflicts and Trade-offs</i>                                        |       |
| Younghun Lee and Dan Goldwasser .....                                                                                                                   | 20848 |
| <i>LogicTree: Structured Proof Exploration for Coherent and Rigorous Logical Reasoning with Large Language Models</i>                                   |       |
| Kang He and Kaushik Roy .....                                                                                                                           | 20864 |
| <i>Unmasking Fake Careers: Detecting Machine-Generated Career Trajectories via Multi-layer Heterogeneous Graphs</i>                                     |       |
| Michiharu Yamashita, Thanh Tran, Delvin Ce Zhang and Dongwon Lee .....                                                                                  | 20894 |
| <i>GAP: a Global Adaptive Pruning Method for Large Language Models</i>                                                                                  |       |
| Zhihua Ban, Haotian Ma, Siheng Zhang, Shengyu Liu, Xichen Chen and Ming Yang .....                                                                      | 20910 |
| <i>Distribution Prompting: Understanding the Expressivity of Language Models Through the Next-Token Distributions They Can Produce</i>                  |       |
| Haojin Wang, Zining Zhu and Freda Shi .....                                                                                                             | 20916 |
| <i>LGA: LLM-GNN Aggregation for Temporal Evolution Attribute Graph Prediction</i>                                                                       |       |
| Feng Zhao, Ruoyu Chai, Kangzheng Liu and Xianggan Liu .....                                                                                             | 20930 |
| <i>EIFBENCH: Extremely Complex Instruction Following Benchmark for Large Language Models</i>                                                            |       |
| Tao Zou, Xinghua Zhang, Haiyang Yu, Minzheng Wang, Fei Huang and Yongbin Li .....                                                                       | 20942 |
| <i>Tool Preferences in Agentic LLMs are Unreliable</i>                                                                                                  |       |
| Kazem Faghih, Wenxiao Wang, Yize Cheng, Siddhant Bharti, Gaurang Sriramanan, Sriram Balasubramanian, Parsa Hosseini and Soheil Feizi .....              | 20966 |
| <i>Enhancing Large Language Model for Knowledge Graph Completion via Structure-Aware Alignment-Tuning</i>                                               |       |
| Yu Liu, Yanan Cao, Xixun Lin, Yanmin Shang, Shi Wang and Shirui Pan .....                                                                               | 20982 |
| <i>MultiDocFusion : Hierarchical and Multimodal Chunking Pipeline for Enhanced RAG on Long Industrial Documents</i>                                     |       |
| Joongmin Shin, Chanjun Park, Jeongbae Park, Jaehyung Seo and Heuseok Lim .....                                                                          | 20997 |



|                                                                                                                                                               |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Attention-guided Self-reflection for Zero-shot Hallucination Detection in Large Language Models</i>                                                        |       |
| Qiang Liu, Xinlong Chen, Yue Ding, Bowen Song, Weiqiang Wang, Shu Wu and Liang Wang                                                                           | 21017 |
| <i>‘Rich Dad, Poor Lad’: How do Large Language Models Contextualize Socioeconomic Factors in College Admission ?</i>                                          |       |
| Huy Nghiem, Phuong-Anh Nguyen-Le, John Prindle, Rachel Rudinger and Hal Daumé Iii                                                                             | 21034 |
| <i>Understanding and Mitigating Overrefusal in LLMs from an Unveiling Perspective of Safety Decision Boundary</i>                                             |       |
| Licheng Pan, Yongqi Tong, Xin Zhang, Xiaolu Zhang, Jun Zhou and Zhixuan Chu . . . . .                                                                         | 21069 |
| <i>MMAG: Multimodal Learning for Mucus Anomaly Grading in Nasal Endoscopy via Semantic Attribute Prompting</i>                                                |       |
| Xinpan Yuan, Mingzhu Huang, Liuji Hua, Jianuo Ju and Xu Zhang . . . . .                                                                                       | 21088 |
| <i>The Emperor’s New Reasoning: Format Imitation Overshadows Genuine Mathematical Understanding in SFT</i>                                                    |       |
| Linyao Yang, Jian-Tao Huang, Yafei Lu, Zhenhui Jessie Li and Guirong Xue . . . . .                                                                            | 21099 |
| <i>Step Guided Reasoning: Improving Mathematical Reasoning using Guidance Generation and Step Reasoning</i>                                                   |       |
| Lang Cao, Yingtian Zou, Chao Peng, Renhong Chen, Wu Ning and Yitong Li . . . . .                                                                              | 21113 |
| <i>Flexibly Utilize Memory for Long-Term Conversation via a Fragment-then-Compose Framework</i>                                                               |       |
| Cai Ke, Yiming Du, Bin Liang, Yifan Xiang, Lin Gui, Zhongyang Li, Baojun Wang, Yue Yu, Hui Wang, Kam-Fai Wong and Ruifeng Xu . . . . .                        | 21131 |
| <i>STRICT: Stress-Test of Rendering Image Containing Text</i>                                                                                                 |       |
| Tianyu Zhang, Xinyu Wang, Lu Li, Zhenghan Tai, Jijun Chi, Jingrui Tian, Hailin He and Suyuchen Wang . . . . .                                                 | 21149 |
| <i>A Sequential Multi-Stage Approach for Code Vulnerability Detection via Confidence- and Collaboration-based Decision Making</i>                             |       |
| Chung-Nan Tsai, Xin Wang, Cheng-Hsiung Lee and Ching-Sheng Lin . . . . .                                                                                      | 21163 |
| <i>Leveraging Large Models to Evaluate Novel Content: A Case Study on Advertisement Creativity</i>                                                            |       |
| Zhaoyi Joey Hou, Adriana Kovashka and Xiang Lorraine Li . . . . .                                                                                             | 21170 |
| <i>BIRD: Bronze Inscription Restoration and Dating</i>                                                                                                        |       |
| Wenjie Hua, Hoang H Nguyen and Gangyan Ge . . . . .                                                                                                           | 21190 |
| <i>DCP: Dual-Cue Pruning for Efficient Large Vision-Language Models</i>                                                                                       |       |
| Lei Jiang, Zixun Zhang, Yuting Zeng, Chunzhao Xie, Tongxuan Liu, Zhen Li, Lechao Cheng and Xiaohua Xu . . . . .                                               | 21203 |
| <i>Improving Context Fidelity via Native Retrieval-Augmented Reasoning</i>                                                                                    |       |
| Suyuchen Wang, Jinlin Wang, Xinyu Wang, Shiqi Li, Xiangru Tang, Sirui Hong, Xiao-Wen Chang, Chenglin Wu and Bang Liu . . . . .                                | 21217 |
| <i>Koel-TTS: Enhancing LLM based Speech Generation with Preference Alignment and Classifier Free Guidance</i>                                                 |       |
| Shehzeen Samarah Hussain, Paarth Neekhara, Xuesong Yang, Edresson Casanova, Subhankar Ghosh, Roy Fejgin, Mikyas T. Desta, Rafael Valle and Jason Li . . . . . | 21231 |
| <i>Mixing Inference-time Experts for Enhancing LLM Reasoning</i>                                                                                              |       |
| Soumya Sanyal, Tianyi Xiao and Xiang Ren . . . . .                                                                                                            | 21247 |

|                                                                                                                                                                                                                                                          |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Reinforced Query Reasoners for Reasoning-intensive Retrieval Tasks</i>                                                                                                                                                                                |       |
| Xubo Qin, Jun Bai, Jiaqi Li, Zixia Jia and Zilong Zheng .....                                                                                                                                                                                            | 21262 |
| <i>TokenSelect: Efficient Long-Context Inference and Length Extrapolation for LLMs via Dynamic Token-Level KV Cache Selection</i>                                                                                                                        |       |
| Wei Wu, Zhuoshi Pan, Kun Fu, Chao Wang, Liyi Chen, Yunchu Bai, Tianfu Wang, Zheng Wang and Hui Xiong .....                                                                                                                                               | 21276 |
| <i>MUSE: MCTS-Driven Red Teaming Framework for Enhanced Multi-Turn Dialogue Safety in Large Language Models</i>                                                                                                                                          |       |
| Siyu Yan, Long Zeng, Xuecheng Wu, Chengcheng Han, Kongcheng Zhang, Chong Peng, Xuezhi Cao, Xunliang Cai and Chenjuan Guo .....                                                                                                                           | 21294 |
| <i>EnAnchored-X2X: English-Anchored Optimization for Many-to-Many Translation</i>                                                                                                                                                                        |       |
| Sen Yang, Yu Bao, Yu Lu, Jiajun Chen, Shujian Huang and Shanbo Cheng .....                                                                                                                                                                               | 21316 |
| <i>“I’ve Decided to Leak”: Probing Internals Behind Prompt Leakage Intents</i>                                                                                                                                                                           |       |
| Jianshuo Dong, Yutong Zhang, Liu Yan, Zhenyu Zhong, Tao Wei, Ke Xu, Minlie Huang, Chao Zhang and Han Qiu .....                                                                                                                                           | 21330 |
| <i>Nullspace Disentanglement for Red Teaming Language Models</i>                                                                                                                                                                                         |       |
| Yi Han, Yuanxing Liu, Weinan Zhang and Ting Liu .....                                                                                                                                                                                                    | 21361 |
| <i>Supervised Attention Mechanism for Low-quality Multimodal Data</i>                                                                                                                                                                                    |       |
| Sijie Mai, Shiqin Han and Haifeng Hu .....                                                                                                                                                                                                               | 21378 |
| <i>Reinforcement Learning for Large Language Models via Group Preference Reward Shaping</i>                                                                                                                                                              |       |
| Huaisheng Zhu, Siyuan Xu, Hangfan Zhang, Teng Xiao, Zhimeng Guo, Shijie Zhou, Shuyue Hu and Vasant G. Honavar .....                                                                                                                                      | 21399 |
| <i>zFLoRA: Zero-Latency Fused Low-Rank Adapters</i>                                                                                                                                                                                                      |       |
| Dhananjaya Gowda, Seoha Song, Harshith Goka and Junhyun Lee .....                                                                                                                                                                                        | 21413 |
| <i>PLAN-TUNING: Post-Training Language Models to Learn Step-by-Step Planning for Complex Problem Solving</i>                                                                                                                                             |       |
| Mihir Parmar, Palash Goyal, Xin Liu, Yiwen Song, Mingyang Ling, Chitta Baral, Hamid Palangi and Tomas Pfister .....                                                                                                                                      | 21431 |
| <i>Semantic Inversion, Identical Replies: Revisiting Negation Blindness in Large Language Models</i>                                                                                                                                                     |       |
| Jinsung Kim, Seonmin Koo and Heuseok Lim .....                                                                                                                                                                                                           | 21446 |
| <i>AMACE: Automatic Multi-Agent Chart Evolution for Iteratively Tailored Chart Generation</i>                                                                                                                                                            |       |
| Hyuk Namgoong, Jeesu Jung, Hyeonseok Kang, Yohan Lee and Sangkeun Jung .....                                                                                                                                                                             | 21484 |
| <i>ActionStudio: A Lightweight Framework for Data and Training of Large Action Models</i>                                                                                                                                                                |       |
| Jianguo Zhang, Thai Quoc Hoang, Ming Zhu, Zuxin Liu, Shiyu Wang, Tulika Manoj Awalgaoonkar, Akshara Prabhakar, Haolin Chen, Weiran Yao, Zhiwei Liu, Juntao Tan, Juan Carlos Niebles, Shelby Heinecke, Huan Wang, Silvio Savarese and Caiming Xiong ..... | 21500 |
| <i>Interpretation Meets Safety: A Survey on Interpretation Methods and Tools for Improving LLM Safety</i>                                                                                                                                                |       |
| Seongmin Lee, Aeree Cho, Grace C. Kim, ShengYun Peng, Mansi Phute and Duen Horng Chau                                                                                                                                                                    | 21515 |
| <i>Unveiling the Response of Large Vision-Language Models to Visually Absent Tokens</i>                                                                                                                                                                  |       |
| Sohee Kim, Soohyun Ryu, Joonhyung Park and Eunho Yang .....                                                                                                                                                                                              | 21547 |



|                                                                                                                                                       |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Improving Task Diversity in Label Efficient Supervised Finetuning of LLMs</i>                                                                      |       |
| Abhinav Arabelly, Jagrut Nemade, Robert D Nowak and Jifan Zhang .....                                                                                 | 21570 |
| <i>Look Beyond Feeling: Unveiling Latent Needs from Implicit Expressions for Proactive Emotional Support</i>                                          |       |
| Xing Fu, Haozhen Li, Bichen Wang, Hao Yang, Yanyan Zhao and Bing Qin .....                                                                            | 21583 |
| <i>s3: You Don't Need That Much Data to Train a Search Agent via RL</i>                                                                               |       |
| Pengcheng Jiang, Xueqiang Xu, Jiacheng Lin, Jinfeng Xiao, Zifeng Wang, Jimeng Sun and Jiawei Han .....                                                | 21611 |
| <i>FuseChat: Knowledge Fusion of Chat Models</i>                                                                                                      |       |
| Fanqi Wan, Longguang Zhong, Ziyi Yang, Ruijun Chen and Xiaojun Quan .....                                                                             | 21630 |
| <i>Continuous-Time Attention: PDE-Guided Mechanisms for Long-Sequence Transformers</i>                                                                |       |
| Yukun Zhang and Xueqing Zhou .....                                                                                                                    | 21655 |
| <i>Forget What You Know about LLMs Evaluations - LLMs are Like a Chameleon</i>                                                                        |       |
| Nurit Cohen Inger, Yehonatan Elisha, Bracha Shapira, Lior Rokach and Seffi Cohen .....                                                                | 21676 |
| <i>Memorization or Reasoning? Exploring the Idiom Understanding of LLMs</i>                                                                           |       |
| Jisu Kim, Youngwoo Shin, Uiji Hwang, Jihun Choi, Richeng Xuan and Taeuk Kim .....                                                                     | 21690 |
| <i>RD-MCSA: A Multi-Class Sentiment Analysis Approach Integrating In-Context Classification Rationales and Demonstrations</i>                         |       |
| Haihua Xie, Yin Zhu Cheng, Yaqing Wang, Miao He and Mingming Sun .....                                                                                | 21712 |
| <i>Puzzled by Puzzles: When Vision-Language Models Can't Take a Hint</i>                                                                              |       |
| Heekyung Lee, Jiaxin Ge, Tsung-Han Wu, Minwoo Kang, Trevor Darrell and David M. Chan                                                                  | 21736 |
| <i>CREPE: Rapid Chest X-ray Report Evaluation by Predicting Multi-category Error Counts</i>                                                           |       |
| Gihun Cho, Seunghyun Jang, Hanbin Ko, Inhyeok Baek and Chang Min Park .....                                                                           | 21750 |
| <i>TIDES: Technical Information Discovery and Extraction System</i>                                                                                   |       |
| Jihee Kim, Subeen Park, Hakyung Lee, YongTaek Lim, Hyo-won Suh and Kyungwoo Song                                                                      | 21768 |
| <i>Learning to Ask: When LLM Agents Meet Unclear Instruction</i>                                                                                      |       |
| Wenxuan Wang, Shi Juluan, Zixuan Ling, Yuk-Kit Chan, Chaozheng Wang, Cheryl Lee, Youliang Yuan, Jen-tse Huang, Wenxiang Jiao and Michael R. Lyu ..... | 21785 |
| <i>RICO: Improving Accuracy and Completeness in Image Recaptioning via Visual Reconstruction</i>                                                      |       |
| Yuchi Wang, Yishuo Cai, Shuhuai Ren, Sihan Yang, Linli Yao, Yuanxin Liu, Yuanxing Zhang, Pengfei Wang and Xu Sun .....                                | 21797 |
| <i>StepSearch: Igniting LLMs Search Ability via Step-Wise Proximal Policy Optimization</i>                                                            |       |
| Xuhui Zheng, Kang An, Ziliang Wang, Yuhang Wang and Yichao Wu .....                                                                                   | 21817 |
| <i>Dynamic Model-Bank Test-Time Adaptation for Automatic Speech Recognition</i>                                                                       |       |
| Yanshuo Wang, Yanghao Zhou, Yukang Lin, Haoxing Chen, Jin Zhang, Wentao Zhu, Jie Hong and Xuesong Li .....                                            | 21843 |
| <i>Mitigating Catastrophic Forgetting in Large Language Models with Forgetting-aware Pruning</i>                                                      |       |
| Wei Huang, Anda Cheng and Yinggui Wang .....                                                                                                          | 21854 |
| <i>Does Localization Inform Unlearning? A Rigorous Examination of Local Parameter Attribution for Knowledge Unlearning in Language Models</i>         |       |
| Hwiyeong Lee, Uiji Hwang, Hyelim Lim and Taeuk Kim .....                                                                                              | 21869 |

|                                                                                                                                                                                                                                                        |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>ArgCMV: An Argument Summarization Benchmark for the LLM-era</i><br>Omkar Gurjar, Agam Goyal and Eshwar Chandrasekharan .....                                                                                                                        | 21882 |
| <i>VistaWise: Building Cost-Effective Agent with Cross-Modal Knowledge Graph for Minecraft</i><br>Honghao Fu, Junlong Ren, Qi Chai, Deheng Ye, Yujun Cai and Hao Wang .....                                                                            | 21896 |
| <i>GraphKV: Breaking the Static Selection Paradigm with Graph-Based KV Cache Eviction</i><br>Xuelin Li, Xiangqi Jin and Linfeng Zhang .....                                                                                                            | 21911 |
| <i>Joint Modeling of Entities and Discourse Relations for Coherence Assessment</i><br>Wei Liu and Michael Strube .....                                                                                                                                 | 21922 |
| <i>Understanding and Leveraging the Expert Specialization of Context Faithfulness in Mixture-of-Experts LLMs</i><br>Jun Bai, Minghao Tong, Yang Liu, Zixia Jia and Zilong Zheng .....                                                                  | 21939 |
| <i>HMoE: Heterogeneous Mixture of Experts for Language Modeling</i><br>An Wang, Xingwu Sun, Ruobing Xie, Shuaipeng Li, Jiaqi Zhu, Zhen Yang, Pinxue Zhao, Weidong Han, Zhanhui Kang, Di Wang, Naoaki Okazaki and Cheng-zhong Xu .....                  | 21955 |
| <i>The Ranking Blind Spot: Decision Hijacking in LLM-based Text Ranking</i><br>Yaoyao Qian, Yifan Zeng, Yuchao Jiang, Chelsi Jain and Huazheng Wang .....                                                                                              | 21970 |
| <i>Uniform Information Density and Syntactic Reduction: Revisiting *that*-Mentioning in English Complement Clauses</i><br>Hailin Hao and Elsi Kaiser .....                                                                                             | 21981 |
| <i>GRIT: Guided Relational Integration for Efficient Multi-Table Understanding</i><br>Yujin Kang, Park Seong Woo and Yoon-Sik Cho .....                                                                                                                | 21996 |
| <i>RPDR: A Round-trip Prediction-Based Data Augmentation Framework for Long-Tail Question Answering</i><br>Yiming Zhang, Siyue Zhang, Junbo Zhao and Chen Zhao .....                                                                                   | 22010 |
| <i>Discrepancy Detection at the Data Level: Toward Consistent Multilingual Question Answering</i><br>Lorena Calvo-Bartolomé, Valérie Aldana, Karla Cantarero, Alonso Madroñal de Mesa, Jerónimo Arenas-García and Jordan Lee Boyd-Graber .....         | 22025 |
| <i>Data-Efficient Selection via Grammatical Complexity in Continual Pre-training of Domain-Specific LLMs</i><br>Yizhou Ying, Geng Zhang, Cui Danxin, Chengyu Du, Guanglei Yue, Sihang Jiang, Jiaqing Liang, Yifei Fu, Hailin Hu and Yanghua Xiao ..... | 22067 |
| <i>Comprehensive and Efficient Distillation for Lightweight Sentiment Analysis Models</i><br>Guangyu Xie, Yice Zhang, Jianzhu Bao, Qianlong Wang, Yang Sun, Bingbing Wang and Ruifeng Xu .....                                                         | 22082 |
| <i>One Planner To Guide Them All ! Learning Adaptive Conversational Planners for Goal-oriented Dialogues</i><br>Huy Quang Dao and Lizi Liao .....                                                                                                      | 22104 |
| <i>Unsupervised Hallucination Detection by Inspecting Reasoning Processes</i><br>Ponhvoan Srey, Xiaobao Wu and Anh Tuan Luu .....                                                                                                                      | 22129 |
| <i>Multimodal Neural Machine Translation: A Survey of the State of the Art</i><br>Yi Feng, Chuanyi Li, Jiatong He, Zhenyu Hou and Vincent Ng .....                                                                                                     | 22142 |

|                                                                                                                                                                                       |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Lemmatization of Polish Multi-word Expressions</i>                                                                                                                                 |       |
| Magdalena Król, Aleksander Smywiński-Pohl, Zbigniew Kaleta and Paweł Lewkowicz . . .                                                                                                  | 22160 |
| <i>Targeted Distillation for Sentiment Analysis</i>                                                                                                                                   |       |
| Yice Zhang, Guangyu Xie, Jingjie Lin, Jianzhu Bao, Qianlong Wang, Xi Zeng and Ruifeng Xu                                                                                              | 22170 |
| <i>DiffusionAttacker: Diffusion-Driven Prompt Manipulation for LLM Jailbreak</i>                                                                                                      |       |
| Hao Wang, Hao Li, Junda Zhu, Xinyuan Wang, Chengwei Pan, Minlie Huang and Lei Sha                                                                                                     | 22194 |
| <i>Rank-Awareness and Angular Constraints: A New Perspective on Learning Sentence Embeddings from NLI Data</i>                                                                        |       |
| Zicheng Zhou, Min Huang and Qinghai Miao . . . . .                                                                                                                                    | 22207 |
| <i>LLM-Guided Semantic Relational Reasoning for Multimodal Intent Recognition</i>                                                                                                     |       |
| Qianrui Zhou, Hua Xu, Yifan Wang, Xinzhi Dong and Hanlei Zhang . . . . .                                                                                                              | 22222 |
| <i>Seeing Culture: A Benchmark for Visual Reasoning and Grounding</i>                                                                                                                 |       |
| Burak Satar, Zhixin Ma, Patrick Amadeus Irawan, Wilfried Ariel Mulyawan, Jing Jiang, Ee-Peng Lim and Chong-Wah Ngo . . . . .                                                          | 22239 |
| <i>GRADA: Graph-based Reranking against Adversarial Documents Attack</i>                                                                                                              |       |
| Jingjie Zheng, Aryo Pradipta Gema, Giwon Hong, Xuanli He, Pasquale Minervini, Youcheng Sun and Qionghai Xu . . . . .                                                                  | 22256 |
| <i>Orchestrating Audio: Multi-Agent Framework for Long-Video Audio Synthesis</i>                                                                                                      |       |
| Yehang Zhang, Xinli Xu, Xiaojie Xu, Doudou Zhang, Li Liu and Ying-Cong Chen . . . . .                                                                                                 | 22279 |
| <i>MADAWSD: Multi-Agent Debate Framework for Adversarial Word Sense Disambiguation</i>                                                                                                |       |
| Kaiyuan Zhang, Qian Liu, Luyang Zhang, Chaoqun Zheng, Shuaimin Li, Bing Xu, Muyun Yang, Xinxiao Qiao and Wenpeng Lu . . . . .                                                         | 22295 |
| <i>Interpretable Text Embeddings and Text Similarity Explanation: A Survey</i>                                                                                                        |       |
| Juri Opitz, Lucas Moeller, Andrianos Michail, Sebastian Padó and Simon Clematide . . . . .                                                                                            | 22315 |
| <i>Dyve: Thinking Fast and Slow for Dynamic Process Verification</i>                                                                                                                  |       |
| Jianyuan Zhong, Zeju Li, Zhijian Xu, Xiangyu Wen and Qiang Xu . . . . .                                                                                                               | 22332 |
| <i>PERSEVAL: A Framework for Perspectivist Classification Evaluation</i>                                                                                                              |       |
| Soda Marem Lo, Silvia Casola, Erhan Sezerer, Valerio Basile, Franco Sansonetti, Antonio Uva and Davide Bernardi . . . . .                                                             | 22346 |
| <i>Massive Supervised Fine-tuning Experiments Reveal How Data, Layer, and Training Factors Shape LLM Alignment Quality</i>                                                            |       |
| Yuto Harada, Yusuke Yamauchi, Yusuke Oda, Yohei Oseki, Yusuke Miyao and Yu Takagi.                                                                                                    | 22372 |
| <i>IndiGEC: Multilingual Grammar Error Correction for Low-Resource Indian Languages</i>                                                                                               |       |
| Ujjwal Sharma and Pushpak Bhattacharyya . . . . .                                                                                                                                     | 22394 |
| <i>Bias Beware: The Impact of Cognitive Biases on LLM-Driven Product Recommendations</i>                                                                                              |       |
| Giorgos Filandrianos, Angeliki Dimitriou, Maria Lymperaio, Konstantinos Thomas and Giorgos Stamou . . . . .                                                                           | 22409 |
| <i>T2R-BENCH: A Benchmark for Real World Table-to-Report Task</i>                                                                                                                     |       |
| Jie Zhang, Changzai Pan, Sishi Xiong, Kaiwen Wei, Yu Zhao, Xiangyu Li, Jiaxin Peng, Xiaoyan Gu, Jian Yang, Wenhan Chang, Zhenhe Wu, Jiang Zhong, Shuangyong Song and Xuelong Li . . . | 22439 |

|                                                                                                                          |       |
|--------------------------------------------------------------------------------------------------------------------------|-------|
| <i>TCP: a Benchmark for Temporal Constraint-Based Planning</i>                                                           |       |
| Zifeng Ding, Sikuan Yan, Moy Yuan, Xianglong Hu, Fangru Lin and Andreas Vlachos . . .                                    | 22464 |
| <i>The Role of Outgoing Connection Heterogeneity in Feedforward Layers of Large Language Models</i>                      |       |
| Felix Stahlberg and Shankar Kumar . . . . .                                                                              | 22488 |
| <i>Follow the Flow: Fine-grained Flowchart Attribution with Neurosymbolic Agents</i>                                     |       |
| Manan Suri, Puneet Mathur, Nedim Lipka, Franck Dernoncourt, Ryan A. Rossi, Vivek Gupta and Dinesh Manocha . . . . .      | 22497 |
| <i>Collaborative Rational Speech Act: Pragmatic Reasoning for Multi-Turn Dialog</i>                                      |       |
| Lautaro Estienne, Gabriel Ben Zenou, Nona Naderi, Jackie CK Cheung and Pablo Piantanida                                  | 22521 |
| <i>Understanding Subword Compositionality of Large Language Models</i>                                                   |       |
| Qiwei Peng, Yekun Chai and Anders Søgaard . . . . .                                                                      | 22536 |
| <i>Internal Chain-of-Thought: Empirical Evidence for Layer-wise Subtask Scheduling in LLMs</i>                           |       |
| Zhipeng Yang, Junzhuo Li, Siyu Xia and Xuming Hu . . . . .                                                               | 22548 |
| <i>From Understanding to Generation: An Efficient Shortcut for Evaluating Language Models</i>                            |       |
| Viktor Hangya, Fabian Küch and Darina Gold . . . . .                                                                     | 22577 |
| <i>Debiasing Multilingual LLMs in Cross-lingual Latent Space</i>                                                         |       |
| Qiwei Peng, Guimin Hu, Yekun Chai and Anders Søgaard . . . . .                                                           | 22594 |
| <i>Context is Gold to find the Gold Passage: Evaluating and Training Contextual Document Embeddings</i>                  |       |
| Max Conti, Manuel Faysse, Gautier Viaud, Antoine Bosselut, Celine Hudelot and Pierre Colombo                             | 22606 |
| <i>MS-RAG: Simple and Effective Multi-Semantic Retrieval-Augmented Generation</i>                                        |       |
| Xiaozhou You, Yahui Luo and Lihong Gu . . . . .                                                                          | 22621 |
| <i>Transitive self-consistency evaluation of NLI models without gold labels</i>                                          |       |
| Wei Wu and Mark Last . . . . .                                                                                           | 22638 |
| <i>MiLQ: Benchmarking IR Models for Bilingual Web Search with Mixed Language Queries</i>                                 |       |
| Jonghwi Kim, Deokhyung Kang, Seonjeong Hwang, Yunsu Kim, Jungseul Ok and Gary Lee                                        | 22655 |
| <i>Enhancing Chinese Offensive Language Detection with Homophonic Perturbation</i>                                       |       |
| Junqi Wu, Shujie Ji, Kang Zhong, Huiling Peng, Zhendongxiao, Xiongding Liu and Wu Wei                                    | 22672 |
| <i>Persona-Augmented Benchmarking: Evaluating LLMs Across Diverse Writing Styles</i>                                     |       |
| Kimberly Truong, Riccardo Fogliato, Hoda Heidari and Steven Wu . . . . .                                                 | 22688 |
| <i>Computational Analysis of Character Development in Holocaust Testimonies</i>                                          |       |
| Esther Shizgal, Eitan Wagner, Renana Keydar and Omri Abend . . . . .                                                     | 22722 |
| <i>TASO: Task-Aligned Sparse Optimization for Parameter-Efficient Model Adaptation</i>                                   |       |
| Daiye Miao, Yufang Liu, Jie Wang, Changzhi Sun, Yunke Zhang, Demei Yan, Shaokang Dong, Qi Zhang and Yuanbin Wu . . . . . | 22747 |
| <i>Dual-Path Counterfactual Integration for Multimodal Aspect-Based Sentiment Classification</i>                         |       |
| Rui Liu, Jiahao Cao, Jiaqian Ren, Xu Bai and Yanan Cao . . . . .                                                         | 22760 |

|                                                                                                                                                                                                                                                                                                                                                                                                        |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Job Unfair: An Investigation of Gender and Occupational Bias in Free-Form Text Completions by LLMs</i><br>Camilla Casula, Sebastiano Vecellio Salto, Elisa Leonardelli and Sara Tonelli .....                                                                                                                                                                                                       | 22771 |
| <i>C3: A Bilingual Benchmark for Spoken Dialogue Models Exploring Challenges in Complex Conversations</i><br>Chengqian Ma, Wei Tao and Steven Y. Guo .....                                                                                                                                                                                                                                             | 22790 |
| <i>Understanding LLMs' Cross-Lingual Context Retrieval: How Good It Is And Where It Comes From</i><br>Changjiang Gao, Hankun Lin, Xin Huang, Xue Han, Junlan Feng, Chao Deng, Jiajun Chen and Shujian Huang .....                                                                                                                                                                                      | 22809 |
| <i>Blind Men and the Elephant: Diverse Perspectives on Gender Stereotypes in Benchmark Datasets</i><br>Mahdi Zakizadeh and Mohammad Taher Pilehvar .....                                                                                                                                                                                                                                               | 22839 |
| <i>Linguistic and Embedding-Based Profiling of Texts Generated by Humans and Large Language Models</i><br>Sergio E. Zanotto and Segun Aroyehun .....                                                                                                                                                                                                                                                   | 22853 |
| <i>An Interdisciplinary Approach to Human-Centered Machine Translation</i><br>Marine Carpuat, Omri Asscher, Kalika Bali, Luisa Bentivogli, Fred Blain, Lynne Bowker, Monojit Choudhury, Hal Daumé Iii, Kevin Duh, Ge Gao, Alvin C Grissom II, Marzena Karpinska, Elaine C Khoong, William D. Lewis, Andre Martins, Mary Nurminen, Douglas W. Oard, Maja Popovic, Michel Simard and François Yvon ..... | 22871 |
| <i>Exploring the Hidden Capacity of LLMs for One-Step Text Generation</i><br>Gleb Mezentsev and Ivan Oseledets .....                                                                                                                                                                                                                                                                                   | 22892 |
| <i>Mixture of Weight-shared Heterogeneous Group Attention Experts for Dynamic Token-wise KV Optimization</i><br>Guanghui Song, Dongping Liao, Yiren Zhao, Kejiang Ye, Cheng-zhong Xu and Xitong Gao                                                                                                                                                                                                    | 22902 |
| <i>PathwiseRAG: Multi-Dimensional Exploration and Integration Framework</i><br>Hengrui Zhang, Pin-Siang Huang, Zhen Zhang, Peican Lin, Yao-Ching Yu, Bo Hu and Yulu Du                                                                                                                                                                                                                                 | 22916 |
| <i>Mm, Wat? Detecting Other-initiated Repair Requests in Dialogue</i><br>Anh Ha Ngo, Nicolas Rollet, Catherine Pelachaud and Chloé Clavel .....                                                                                                                                                                                                                                                        | 22938 |
| <i>R-BPE: Improving BPE-Tokenizers with Token Reuse</i><br>Nancy Hamdan, Osama Rakan Al Mraikhat and Fadi Zaraket .....                                                                                                                                                                                                                                                                                | 22952 |
| <i>Language Models Can be Efficiently Steered via Minimal Embedding Layer Transformations</i><br>Diogo Tavares, David Semedo, Alexander Rudnicky and Joao Magalhaes .....                                                                                                                                                                                                                              | 22961 |
| <i>Adversarial Attacks Against Automated Fact-Checking: A Survey</i><br>Fanzhen Liu, Sharif Abuadbba, Kristen Moore, Surya Nepal, Cecile Paris, Jia Wu, Jian Yang and Quan Z. Sheng .....                                                                                                                                                                                                              | 22980 |
| <i>WildDoc: How Far Are We from Achieving Comprehensive and Robust Document Understanding in the Wild?</i><br>An-Lan Wang, Jingqun Tang, Lei Liao, Hao Feng, Qi Liu, Xiang Fei, Jinghui Lu, Han Wang, Hao Liu, Yuliang Liu, Xiang Bai and Can Huang .....                                                                                                                                              | 23003 |
| <i>DCR: Quantifying Data Contamination in LLMs Evaluation</i><br>Cheng Xu, Nan Yan, Shuhao Guan, Changhong Jin, Yuke Mei, Yibing Guo and Tahar Kechadi                                                                                                                                                                                                                                                 | 23014 |

|                                                                                                                                                                                             |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Building Trust in Clinical LLMs: Bias Analysis and Dataset Transparency</i>                                                                                                              |       |
| Svetlana Maslenskova, Clement Christophe, Marco AF Pimentel, Tathagata Raha, Muhammad Umar Salman, Ahmed Al Mahrooqi, Avani Gupta, Shadab Khan, Ronnie Rajan and Praveenkumar Kanithi ..... | 23033 |
| <i>Surprise Calibration for Better In-Context Learning</i>                                                                                                                                  |       |
| Zhihang Tan, Jingrui Hou, Ping Wang, Qibiao Hu and Peng Zhu .....                                                                                                                           | 23057 |
| <i>SPARK: Simulating the Co-evolution of Stance and Topic Dynamics in Online Discourse with LLM-based Agents</i>                                                                            |       |
| Bowen Zhang, Yi Yang, Fuqiang Niu, Xianghua Fu, Genan Dai and Hu Huang .....                                                                                                                | 23073 |
| <i>Drivel-ology: Challenging LLMs with Interpreting Nonsense with Depth</i>                                                                                                                 |       |
| Yang Wang, Chenghao Xiao, Chia-Yi Hsiao, Zi Yan Chang, Chi-Li Chen, Tyler Loakman and Chenghua Lin .....                                                                                    | 23086 |
| <i>Can Large Language Models be Effective Online Opinion Miners?</i>                                                                                                                        |       |
| Ryang Heo, Yongsik Seo, Junseong Lee and Dongha Lee .....                                                                                                                                   | 23109 |
| <i>Can Large Language Models Translate Unseen Languages in Underrepresented Scripts?</i>                                                                                                    |       |
| Dianqing Lin, Aruukhan, Hongxu Hou, Shuo Sun, Wei Chen, Yichen Yang and Guo Dong Shi                                                                                                        | 23149 |
| <i>InterIDEAS: Philosophical Intertextuality via LLMs</i>                                                                                                                                   |       |
| Yue Yang, Yinzi Xu, Chenghao Huang, JohnMichael Jurgensen, Han Hu and Hao Wang .                                                                                                            | 23163 |
| <i>KCS: Diversify Multi-hop Question Generation with Knowledge Composition Sampling</i>                                                                                                     |       |
| Yangfan Wang, Jie Liu, Chen Tang, Lian Yan and Jingchi Jiang .....                                                                                                                          | 23185 |
| <i>Fooling the LVLMM Judges: Visual Biases in LVLMM-Based Evaluation</i>                                                                                                                    |       |
| Yerin Hwang, Dongryeol Lee, Kyungmin Min, Taegwan Kang, Yongil Kim and Kyomin Jung                                                                                                          | 23198 |
| <i>Disentangled Information Bottleneck for Adversarial Text Defense</i>                                                                                                                     |       |
| Yidan Xu, Xinghao Yang, Wei Liu, Bao-di Liu and Weifeng Liu .....                                                                                                                           | 23218 |
| <i>How do Language Models Reshape Entity Alignment? A Survey of LM-Driven EA Methods: Advances, Benchmarks, and Future</i>                                                                  |       |
| Zerui Chen, Huiming Fan, Qianyu Wang, Tao He, Ming Liu, Heng Chang, Weijiang Yu, Ze Li and Bing Qin .....                                                                                   | 23231 |
| <i>Enhancing LLM-Based Social Bot via an Adversarial Learning Framework</i>                                                                                                                 |       |
| Fanqi Kong, Xiaoyuan Zhang, Xinyu Chen, Yaodong Yang, Song-Chun Zhu and Xue Feng                                                                                                            | 23247 |
| <i>GER-LLM: Efficient and Effective Geospatial Entity Resolution with Large Language Model</i>                                                                                              |       |
| Haojia Zhu, Zhicheng Li and Jiahui Jin .....                                                                                                                                                | 23273 |
| <i>CodeRAG: Finding Relevant and Necessary Knowledge for Retrieval-Augmented Repository-Level Code Completion</i>                                                                           |       |
| Sheng Zhang, Yifan Ding, Shuquan Lian, Shun Song and Hui Li .....                                                                                                                           | 23290 |
| <i>Searching for the Most Human-like Emergent Language</i>                                                                                                                                  |       |
| Brendon Boldt and David R. Mortensen .....                                                                                                                                                  | 23301 |
| <i>Does Context Matter? A Prosodic Comparison of English and Spanish in Monolingual and Multilingual Discourse Settings</i>                                                                 |       |
| Debasmita Bhattacharya, David Sasu, Michela Marchini, Natalie Schluter and Julia Hirschberg                                                                                                 | 23320 |



|                                                                                                                                              |       |
|----------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>ZERA: Zero-init Instruction Evolving Refinement Agent – From Zero Instructions to Structured Prompts via Principle-based Optimization</i> |       |
| Seungyoun Yi, Minsoo Khang and Sungrae Park . . . . .                                                                                        | 23335 |
| <i>Toward Machine Interpreting: Lessons from Human Interpreting Studies</i>                                                                  |       |
| Matthias Sperber, Maureen de Seyssel, Jiajun Bao and Matthias Paulik . . . . .                                                               | 23350 |
| <i>FlashAdventure: A Benchmark for GUI Agents Solving Full Story Arcs in Diverse Adventure Games</i>                                         |       |
| Jaewoo Ahn, Junseo Kim, Heeseung Yun, Jaehyeon Son, Dongmin Park, Jaewoong Cho and Gunhee Kim . . . . .                                      | 23366 |
| <i>FLARE: Faithful Logic-Aided Reasoning and Exploration</i>                                                                                 |       |
| Erik Arakelyan, Pasquale Minervini, Patrick Lewis, Pat Verga and Isabelle Augenstein . . . .                                                 | 23397 |
| <i>Discourse-Driven Code-Switching: Analyzing the Role of Content and Communicative Function in Spanish-English Bilingual Speech</i>         |       |
| Debasmita Bhattacharya, Juan Junco, Divya Tadimeti and Julia Hirschberg . . . . .                                                            | 23416 |
| <i>Can Large Language Models Translate Spoken-Only Languages through International Phonetic Transcription?</i>                               |       |
| Jiale Chen, Xuelian Dong, Qihao Yang, Wenxiu Xie and Tianyong Hao . . . . .                                                                  | 23432 |
| <i>ClimateViz: A Benchmark for Statistical Reasoning and Fact Verification on Scientific Charts</i>                                          |       |
| Ruiran Su, Jiasheng Si, Zhijiang Guo and Janet B. Pierrehumbert . . . . .                                                                    | 23448 |
| <i>Bridging the Gap Between Molecule and Textual Descriptions via Substructure-aware Alignment</i>                                           |       |
| Hyuntae Park, Yeachan Kim and SangKeun Lee . . . . .                                                                                         | 23471 |
| <i>SLlama: Parameter-Efficient Language Model Architecture for Enhanced Linguistic Competence Under Strict Data Constraints</i>              |       |
| Victor Adelakun Omolaoye, Babajide Alamu Owoyele and Gerard de Melo . . . . .                                                                | 23492 |
| <i>What You See is What You Ask: Evaluating Audio Descriptions</i>                                                                           |       |
| Divy Kala, Eshika Khandelwal and Makarand Tapaswi . . . . .                                                                                  | 23508 |
| <i>TAPS: Tool-Augmented Personalisation via Structured Tagging</i>                                                                           |       |
| Ekaterina Taktasheva and Jeff Dalton . . . . .                                                                                               | 23531 |
| <i>Investigating How Pre-training Data Leakage Affects Models' Reproduction and Detection Capabilities</i>                                   |       |
| Masahiro Kaneko and Timothy Baldwin . . . . .                                                                                                | 23557 |
| <i>Walk and Read Less: Improving the Efficiency of Vision-and-Language Navigation via Tuning-Free Multimodal Token Pruning</i>               |       |
| Wenda Qin, Andrea Burns, Bryan A. Plummer and Margrit Betke . . . . .                                                                        | 23568 |
| <i>Connecting the Knowledge Dots: Retrieval-augmented Knowledge Connection for Commonsense Reasoning</i>                                     |       |
| Junho Kim, Soyeon Bak, Mingyu Lee, Minju Hong, Songha Kim, Tae-Eui Kam and SangKeun Lee . . . . .                                            | 23583 |
| <i>Agent-as-Judge for Factual Summarization of Long Narratives</i>                                                                           |       |
| Yeonseok Jeong, Minsoo Kim, Seung-won Hwang and Byung-Hak Kim . . . . .                                                                      | 23603 |
| <i>DnDScore: Decontextualization and Decomposition for Factuality Verification in Long-Form Text Generation</i>                              |       |
| Miriam Wanner, Benjamin Van Durme and Mark Dredze . . . . .                                                                                  | 23621 |

|                                                                                                                                                                                                                                                                            |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>RAcQUeT: Unveiling the Dangers of Overlooked Referential Ambiguity in Visual LLMs</i>                                                                                                                                                                                   |       |
| Alberto Testoni, Barbara Plank and Raquel Fernández .....                                                                                                                                                                                                                  | 23639 |
| <i>Resource-Rational Noisy-Channel Language Processing: Testing the Effect of Algorithmic Constraints on Inferences</i>                                                                                                                                                    |       |
| Thomas Hikaru Clark, Jacob Hoover Vigly, Edward Gibson and Roger P. Levy .....                                                                                                                                                                                             | 23660 |
| <i>In Benchmarks We Trust ... Or Not?</i>                                                                                                                                                                                                                                  |       |
| Ine Gevers, Victor De Marez, Jens Van Nooten, Jens Lemmens, Andriy Kosar, Ehsan Lotfi, Nikolay Banar, Pieter Fivez, Luna De Bruyne and Walter Daelemans .....                                                                                                              | 23674 |
| <i>Video2Roleplay: A Multimodal Dataset and Framework for Video-Guided Role-playing Agents</i>                                                                                                                                                                             |       |
| Xueqiao Zhang, Chao Zhang, Jingtao Xu, Yifan Zhu, Xin Shi, Yi Yang and Yawei Luo ...                                                                                                                                                                                       | 23689 |
| <i>Discriminating Form and Meaning in Multilingual Models with Minimal-Pair ABX Tasks</i>                                                                                                                                                                                  |       |
| Maureen de Seyssel, Jie Chi, Skyler Seto, Maartje Ter Hoeve, Masha Fedzechkina and Natalie Schluter .....                                                                                                                                                                  | 23716 |
| <i>Rethinking Text-based Protein Understanding: Retrieval or LLM?</i>                                                                                                                                                                                                      |       |
| Juntong Wu, Zijing Liu, He Cao, Li Hao, Bin Feng, Zishan Shu, Ke Yu, Li Yuan and Yu Li                                                                                                                                                                                     | 23738 |
| <i>Grounded Semantic Role Labelling from Synthetic Multimodal Data for Situated Robot Commands</i>                                                                                                                                                                         |       |
| Claudiu Daniel Hromei, Antonio Scaiella, Danilo Croce and Roberto Basili .....                                                                                                                                                                                             | 23759 |
| <i>Easy as PIE? Identifying Multi-Word Expressions with LLMs</i>                                                                                                                                                                                                           |       |
| Kai Golan Hashiloni, Ofri Hefetz and Kfir Bar .....                                                                                                                                                                                                                        | 23783 |
| <i>Query-Focused Retrieval Heads Improve Long-Context Reasoning and Re-ranking</i>                                                                                                                                                                                         |       |
| Wuwei Zhang, Fangcong Yin, Howard Yen, Danqi Chen and Xi Ye .....                                                                                                                                                                                                          | 23803 |
| <i>Robust Adaptation of Large Multimodal Models for Retrieval Augmented Hateful Meme Detection</i>                                                                                                                                                                         |       |
| Jingbiao Mei, Jinghong Chen, Guangyu Yang, Weizhe Lin and Bill Byrne .....                                                                                                                                                                                                 | 23818 |
| <i>Audio-Reasoner: Improving Reasoning Capability in Large Audio Language Models</i>                                                                                                                                                                                       |       |
| Xie Zhifei, Mingbao Lin, Zihang Liu, Pengcheng Wu, Shuicheng Yan and Chunyan Miao .                                                                                                                                                                                        | 23841 |
| <i>From perception to production: how acoustic invariance facilitates articulatory learning in a self-supervised vocal imitation model</i>                                                                                                                                 |       |
| Marvin Lavechin and Thomas Hueber .....                                                                                                                                                                                                                                    | 23864 |
| <i>REALM: Recursive Relevance Modeling for LLM-based Document Re-Ranking</i>                                                                                                                                                                                               |       |
| Pinhuan Wang, Zhiqiu Xia, Chunhua Liao, Feiyi Wang and Hang Liu .....                                                                                                                                                                                                      | 23876 |
| <i>PLLuM-Align: Polish Preference Dataset for Large Language Model Alignment</i>                                                                                                                                                                                           |       |
| Karolina Seweryn, Anna Kołos, Agnieszka Karlińska, Katarzyna Lorenc, Katarzyna Dziejulska, Maciej Chrabaszcz, Aleksandra Krasnodebska, Paula Betscher, Zofia Cieślińska, Katarzyna Kowol, Julia Moska, Dawid Motyka, Paweł Walkowiak, Bartosz Żuk and Arkadiusz Janz ..... | 23891 |
| <i>Graph-R1: Incentivizing the Zero-Shot Graph Learning Capability in LLMs via Explicit Reasoning</i>                                                                                                                                                                      |       |
| Yicong Wu, Guangyue Lu, Yuan Zuo, Huarong Zhang and Junjie Wu .....                                                                                                                                                                                                        | 23921 |
| <i>Scalable and Culturally Specific Stereotype Dataset Construction via Human-LLM Collaboration</i>                                                                                                                                                                        |       |
| Weicheng Ma, John J. Guerrerio and Soroush Vosoughi .....                                                                                                                                                                                                                  | 23940 |
| <i>Can Large Language Models Be Good Language Teachers?</i>                                                                                                                                                                                                                |       |
| LiQing Xu, Qiwei Li, Tianshuo Peng, Zuchao Li, Hai Zhao and Ping Wang .....                                                                                                                                                                                                | 23969 |



|                                                                                                                                                                            |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Empowering Math Problem Generation and Reasoning for Large Language Model via Synthetic Data based Continual Learning Framework</i>                                     |       |
| Qian Wan, Wangzi Shi, Jintian Feng, Shengyingjie Liu, Luona Wei, Zhicheng Dai and Jianwen Sun                                                                              | 23984 |
| <i>Tokenization and Representation Biases in Multilingual Models on Dialectal NLP Tasks</i>                                                                                |       |
| Vani Kanjirangat, Tanja Samardzic, Ljiljana Dolamic and Fabio Rinaldi                                                                                                      | 24004 |
| <i>Evaluating the Evaluators: Are readability metrics good measures of readability?</i>                                                                                    |       |
| Isabel Cachola, Daniel Khashabi and Mark Dredze                                                                                                                            | 24023 |
| <i>Text Takes Over: A Study of Modality Bias in Multimodal Intent Detection</i>                                                                                            |       |
| Ankan Mullick, Saransh Sharma, Abhik Jana and Pawan Goyal                                                                                                                  | 24040 |
| <i>What's in a prompt? Language models encode literary style in prompt embeddings</i>                                                                                      |       |
| Raphaël Sarfati, Haley Moller, Toni J.b. Liu, Nicolas Boulle and Christopher Earls                                                                                         | 24071 |
| <i>Identifying and Answering Questions with False Assumptions: An Interpretable Approach</i>                                                                               |       |
| Zijie Wang and Eduardo Blanco                                                                                                                                              | 24081 |
| <i>VisFinEval: A Scenario-Driven Chinese Multimodal Benchmark for Holistic Financial Understanding</i>                                                                     |       |
| Zhaowei Liu, Xin Guo, Haotian Xia, Lingfeng Zeng, Fangqi Lou, Jinyi Niu, Mengping Li, Qi Qi, Jiahuan Li, Wei Zhang, Yinglong Wang, Weige Cai, Weining Shen and Liwen Zhang | 24100 |
| <i>Socratic-MCTS: Test-Time Visual Reasoning by Asking the Right Questions</i>                                                                                             |       |
| David Acuna, Ximing Lu, Jaehun Jung, Hyunwoo Kim, Amlan Kar, Sanja Fidler and Yejin Choi                                                                                   | 24159 |
| <i>LLMs Don't Know Their Own Decision Boundaries: The Unreliability of Self-Generated Counterfactual Explanations</i>                                                      |       |
| Harry Mayne, Ryan Othniel Kearns, Yushi Yang, Andrew M. Bean, Eoin D. Delaney, Chris Russell and Adam Mahdi                                                                | 24173 |
| <i>Grounding Multilingual Multimodal LLMs With Cultural Knowledge</i>                                                                                                      |       |
| Jean De Dieu Nyandwi, Yueqi Song, Simran Khanuja and Graham Neubig                                                                                                         | 24199 |
| <i>Following Length Constraints in Instructions</i>                                                                                                                        |       |
| Weizhe Yuan, Ilia Kulikov, Ping Yu, Kyunghyun Cho, Sainbayar Sukhbaatar, Jason E Weston and Jing Xu                                                                        | 24244 |
| <i>Memory-QA: Answering Recall Questions Based on Multimodal Memories</i>                                                                                                  |       |
| Hongda Jiang, Xinyuan Zhang, Siddhant Garg, Rishab Arora, Shiun-Zu Kuo, Jiayang Xu, Aaron Colak and Xin Luna Dong                                                          | 24256 |
| <i>NEXUS: Network Exploration for eXploiting Unsafe Sequences in Multi-Turn LLM Jailbreaks</i>                                                                             |       |
| Javad Rafiei Asl, Sidhant Narula, Mohammad Ghasemigol, Eduardo Blanco and Daniel Takabi                                                                                    | 24279 |
| <i>Sketch-of-Thought: Efficient LLM Reasoning with Adaptive Cognitive-Inspired Sketching</i>                                                                               |       |
| Simon A. Aytes, Jinheon Baek and Sung Ju Hwang                                                                                                                             | 24308 |
| <i>From Language to Cognition: How LLMs Outgrow the Human Language Network</i>                                                                                             |       |
| Badr AlKhamissi, Greta Tuckute, Yingtian Tang, Taha Osama A Bihuraib, Antoine Bosselut and Martin Schrimpf                                                                 | 24333 |
| <i>Logos as a Well-Tempered Pre-train for Sign Language Recognition</i>                                                                                                    |       |
| Ilya Ovodov, Petr Surovtsev, Karina Kvanchiani, Alexander Kapitanov and Alexander Nagaev                                                                                   | 24352 |

|                                                                                                                                                       |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Hallucination Detection in LLMs Using Spectral Features of Attention Maps</i>                                                                      |       |
| Jakub Binkowski, Denis Janiak, Albert Sawczyn, Bogdan Gabrys and Tomasz Jan Kajdanowicz                                                               | 24366 |
| <i>Composable Cross-prompt Essay Scoring by Merging Models</i>                                                                                        |       |
| Sanwoo Lee, Kun Liang and Yunfang Wu                                                                                                                  | 24398 |
| <i>Towards a Holistic and Automated Evaluation Framework for Multi-Level Comprehension of LLMs in Book-Length Contexts</i>                            |       |
| Yuhoo Lee, Jiaqi Deng, Nicole Hee-Yeon Kim, Hyangsuk Min, Taewon Yun, Minjeong Ban, Kim Yul and Hwanjun Song                                          | 24413 |
| <i>Improving Large Language Models Function Calling and Interpretability via Guided-Structured Templates</i>                                          |       |
| Hy Dang, Tianyi Liu, Zhuofeng Wu, Jingfeng Yang, Haoming Jiang, Tao Yang, Pei Chen, Zhen-gyang Wang, Helen Wang, Huasheng Li, Bing Yin and Meng Jiang | 24438 |
| <i>Evaluation and Facilitation of Online Discussions in the LLM Era: A Survey</i>                                                                     |       |
| Katerina Korre, Dimitris Tsirmpas, Nikos Gkoumas, Emma Cabalé, Danai Myrtzani, Theodoros Evgeniou, Ion Androutsopoulos and John Pavlopoulos           | 24455 |
| <i>Temporal Scaling Law for Large Language Models</i>                                                                                                 |       |
| Yizhe Xiong, Xiansheng Chen, Xin Ye, Hui Chen, Zijia Lin, Haoran Lian, Zhenpeng Su, Wei Huang, Jianwei Niu, Jungong Han and Guiguang Ding             | 24475 |
| <i>Reframe Your Life Story: Interactive Narrative Therapist and Innovative Moment Assessment with Large Language Models</i>                           |       |
| Yi Feng, Jiaqi Wang, Wenxuan Zhang, Zhuang Chen, Shen Yutong, Xiyao Xiao, Minlie Huang, Liping Jing and Jian Yu                                       | 24496 |
| <i>From Word to World: Evaluate and Mitigate Culture Bias in LLMs via Word Association Test</i>                                                       |       |
| Xunlian Dai, Li Zhou, Benyou Wang and Haizhou Li                                                                                                      | 24522 |
| <i>Mitigating the Privacy Issues in Retrieval-Augmented Generation (RAG) via Pure Synthetic Data</i>                                                  |       |
| Shenglai Zeng, Jiankun Zhang, Pengfei He, Jie Ren, Tianqi Zheng, Hanqing Lu, Han Xu, Hui Liu, Yue Xing and Jiliang Tang                               | 24539 |
| <i>AdaSteer: Your Aligned LLM is Inherently an Adaptive Jailbreak Defender</i>                                                                        |       |
| Weixiang Zhao, Jiahe Guo, Yulin Hu, Yang Deng, An Zhang, Xingyu Sui, Xinyang Han, Yanyan Zhao, Bing Qin, Tat-Seng Chua and Ting Liu                   | 24571 |
| <i>Large Language Models Meet Knowledge Graphs for Question Answering: Synthesis and Opportunities</i>                                                |       |
| Chuangtao Ma, Yongrui Chen, Tianxing Wu, Arijit Khan and Haofen Wang                                                                                  | 24590 |
| <i>TFDP: Token-Efficient Disparity Audits for Autoregressive LLMs via Single-Token Masked Evaluation</i>                                              |       |
| Inderjeet Singh, Ramya Srinivasan, Roman Vainshtein and Hisashi Kojima                                                                                | 24610 |
| <i>Hanfu-Bench: A Multimodal Benchmark on Cross-Temporal Cultural Understanding and Transcreation</i>                                                 |       |
| Li Zhou, Lutong Yu, Dongchu Xie, Shaohuan Cheng, Wenyan Li and Haizhou Li                                                                             | 24628 |
| <i>MERMAID: Multi-perspective Self-reflective Agents with Generative Augmentation for Emotion Recognition</i>                                         |       |
| Zhongyu Yang, Junhao Song, Siyang Song, Wei Pang and Yingfang Yuan                                                                                    | 24651 |
| <i>Personality Vector: Modulating Personality of Large Language Models by Model Merging</i>                                                           |       |
| Seungjong Sun, Seo Yeon Baek and Jang Hyun Kim                                                                                                        | 24668 |

|                                                                                                                                                    |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Beyond Outlining: Heterogeneous Recursive Planning for Adaptive Long-form Writing with Language Models</i>                                      |       |
| Ruibin Xiong, Yimeng Chen, Dmitrii Khizbullin, Mingchen Zhuge and Jürgen Schmidhuber                                                               | 24690 |
| <i>Hidden in Plain Sight: Reasoning in Underspecified and Misspecified Scenarios for Multimodal LLMs</i>                                           |       |
| Qianqi Yan, Hongquan Li, Shan Jiang, Yang Zhao, Xinze Guan, Ching-Chen Kuo and Xin Eric Wang                                                       | 24727 |
| <i>PrimeX: A Dataset of Worldview, Opinion, and Explanation</i>                                                                                    |       |
| Rik Koncel-Kedziorski, Brihi Joshi and Tim Paek                                                                                                    | 24748 |
| <i>LASER: An LLM-based ASR Scoring and Evaluation Rubric</i>                                                                                       |       |
| Amruta Parulekar and Preethi Jyothi                                                                                                                | 24774 |
| <i>Improving Zero-shot Sentence Decontextualisation with Content Selection and Planning</i>                                                        |       |
| Zhenyun Deng, Yulong Chen and Andreas Vlachos                                                                                                      | 24784 |
| <i>Beyond Text: Unveiling Privacy Vulnerabilities in Multi-modal Retrieval-Augmented Generation</i>                                                |       |
| Jiankun Zhang, Shenglai Zeng, Jie Ren, Tianqi Zheng, Hui Liu, Xianfeng Tang, Hui Liu and Yi Chang                                                  | 24801 |
| <i>Code Execution as Grounded Supervision for LLM Reasoning</i>                                                                                    |       |
| Dongwon Jung, Wenxuan Zhou and Muhao Chen                                                                                                          | 24823 |
| <i>Subjective Behaviors and Preferences in LLM: Language of Browsing</i>                                                                           |       |
| Sai Sundaresan, Harshita Chopra, Atanu R. Sinha, Koustava Goswami, Nagasai Saketh Naidu, Raghav Karan and N Anushka                                | 24835 |
| <i>Pixels Versus Priors: Controlling Knowledge Priors in Vision-Language Models through Visual Counterfactuals</i>                                 |       |
| Michal Golovanevsky, William Rudman, Michael A. Lepori, Amir Bar, Ritambhara Singh and Carsten Eickhoff                                            | 24849 |
| <i>Balcony: A Lightweight Approach to Dynamic Inference of Generative Language Models</i>                                                          |       |
| Benyamin Jamialahmadi, Parsa Kavehzadeh, Mehdi Rezagholizadeh, Parsa Farinneya, Hossein Rajabzadeh, Aref Jafari, Boxing Chen and Marzieh S. Tahaei | 24865 |
| <i>Social Genome: Grounded Social Reasoning Abilities of Multimodal Models</i>                                                                     |       |
| Leena Mathur, Marian Qian, Paul Pu Liang and Louis-Philippe Morency                                                                                | 24880 |
| <i>Profiler: Black-box AI-generated Text Origin Detection via Context-aware Inference Pattern Analysis</i>                                         |       |
| Hanxi Guo, Siyuan Cheng, Xiaolong Jin, Zhuo Zhang, Guangyu Shen, Kaiyuan Zhang, Shengwei An, Guan hong Tao and Xiangyu Zhang                       | 24904 |
| <i>Speech Discrete Tokens or Continuous Features? A Comparative Analysis for Spoken Language Understanding in SpeechLLMs</i>                       |       |
| Dingdong Wang, Junan Li, Mingyu Cui, Dongchao Yang, Xueyuan Chen and Helen M. Meng                                                                 | 24925 |
| <i>RAG-Zeval: Enhancing RAG Responses Evaluator through End-to-End Reasoning and Ranking-Based Reinforcement Learning</i>                          |       |
| Kun LI, Yunxiang Li, Tianhua Zhang, Hongyin Luo, Xixin Wu, James R. Glass and Helen M. Meng                                                        | 24937 |
| <i>Infini-gram mini: Exact n-gram Search at the Internet Scale with FM-Index</i>                                                                   |       |
| Hao Xu, Jiacheng Liu, Yejin Choi, Noah A. Smith and Hannaneh Hajishirzi                                                                            | 24956 |

|                                                                                                                                                                                                                   |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Mahānāma: A Unique Testbed for Literary Entity Discovery and Linking</i>                                                                                                                                       |       |
| Sujoy Sarkar, Gourav Sarkar, Manoj Balaji Jagadeeshan, Jivnesh Sandhan, Amrith Krishna and Pawan Goyal . . . . .                                                                                                  | 24982 |
| <i>Adaptively profiling models with task elicitation</i>                                                                                                                                                          |       |
| Davis Brown, Prithvi Balehannina, Helen Jin, Shreya Havaladar, Hamed Hassani and Eric Wong                                                                                                                        | 24997 |
| <i>Causal Interventions Reveal Shared Structure Across English Filler–Gap Constructions</i>                                                                                                                       |       |
| Sasha Boguraev, Christopher Potts and Kyle Mahowald . . . . .                                                                                                                                                     | 25033 |
| <i>TactfulToM: Do LLMs have the Theory of Mind ability to understand White Lies?</i>                                                                                                                              |       |
| Yiwei Liu, Emma Jane Pretty, Jiahao Huang and Saku Sugawara . . . . .                                                                                                                                             | 25055 |
| <i>Don’t Sweat the Small Stuff: Segment-Level Meta-Evaluation Based on Pairwise Difference Correlation</i>                                                                                                        |       |
| Colten DiIanni and Daniel Deutsch . . . . .                                                                                                                                                                       | 25074 |
| <i>SMART: Simulated Students Aligned with Item Response Theory for Question Difficulty Prediction</i>                                                                                                             |       |
| Alexander Scarlatos, Nigel Fernandez, Christopher Ormerod, Susan Lottridge and Andrew Lan                                                                                                                         | 25083 |
| <i>HESEIA: A community-based dataset for evaluating social biases in large language models, co-designed in real school settings in Latin America</i>                                                              |       |
| Guido Ivetta, Marcos J Gomez, Sofia Martinelli, Pietro Palombini, M Emilia Echeveste, Nair Carolina Mazzeo, Beatriz Busaniche and Luciana Benotti . . . . .                                                       | 25107 |
| <i>WebMMU: A Benchmark for Multimodal Multilingual Website Understanding and Code Generation</i>                                                                                                                  |       |
| Rabiul Awal, Mahsa Massoud, Aarash Feizi, Zichao Li, Suyuchen Wang, Christopher Pal, Aishwarya Agrawal, David Vazquez, Siva Reddy, Juan A. Rodriguez, Perouz Taslakian, Spandana Gella and Sai Rajeswar . . . . . | 25130 |
| <i>Analyzing values about gendered language reform in LLMs’ revisions</i>                                                                                                                                         |       |
| Jules Watson, Xi Wang, Raymond Liu, Suzanne Stevenson and Barend Beekhuizen . . . . .                                                                                                                             | 25158 |
| <i>ALLabel: Three-stage Active Learning for LLM-based Entity Recognition using Demonstration Retrieval</i>                                                                                                        |       |
| Zihan Chen, Lei Shi, Weize Wu, Qiji Zhou and Yue Zhang . . . . .                                                                                                                                                  | 25174 |
| <i>HyperKGR: Knowledge Graph Reasoning in Hyperbolic Space with Graph Neural Network Encoding Symbolic Path</i>                                                                                                   |       |
| Lihui Liu . . . . .                                                                                                                                                                                               | 25189 |
| <i>LLaMP: Large Language Model Made Powerful for High-fidelity Materials Knowledge Retrieval</i>                                                                                                                  |       |
| Yuan Chiang, Elvis Hsieh, Chia-Hong Chou and Janosh Riebesell . . . . .                                                                                                                                           | 25201 |
| <i>ReSeeding Latent States for Sequential Language Understanding</i>                                                                                                                                              |       |
| Stéphane Aroca-Ouellette, Katharina von der Wense and Alessandro Roncone . . . . .                                                                                                                                | 25234 |
| <i>DPED: Multi-Layer Noise Distillation for Privacy-Preserving Text Embeddings</i>                                                                                                                                |       |
| Shuya Feng and Yuan Hong . . . . .                                                                                                                                                                                | 25249 |
| <i>Identifying &amp; Interactively Refining Ambiguous User Goals for Data Visualization Code Generation</i>                                                                                                       |       |
| Mert Inan, Anthony Sicilia, Alex Xie, Saujas Vaduguru, Daniel Fried and Malihe Alikhani                                                                                                                           | 25258 |
| <i>Morpheme Induction for Emergent Language</i>                                                                                                                                                                   |       |
| Brendon Boldt and David R. Mortensen . . . . .                                                                                                                                                                    | 25276 |

|                                                                                                                                                          |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Stepwise Informativeness Search for Improving LLM Reasoning</i>                                                                                       |       |
| Siyuan Wang, Enda Zhao and Xiang Ren .....                                                                                                               | 25292 |
| <i>Social Good or Scientific Curiosity? Uncovering the Research Framing Behind NLP Artefacts</i>                                                         |       |
| Eric Chamoun, Nedjma Ousidhoum, Michael Sejr Schlichtkrull and Andreas Vlachos ....                                                                      | 25311 |
| <i>FairGen: Controlling Sensitive Attributes for Fair Generations in Diffusion Models via Adaptive Latent Guidance</i>                                   |       |
| Mintong Kang, Vinayshekhar Bannihatti Kumar, Shamik Roy, Abhishek Kumar, Sopan Khosla, Balakrishnan Murali Narayanaswamy and Rashmi Gangadharaiyah ..... | 25348 |
| <i>Contra4: Evaluating Contrastive Cross-Modal Reasoning in Audio, Video, Image, and 3D</i>                                                              |       |
| Artemis Panagopoulou, Le Xue, Honglu Zhou, Silvio Savarese, Ran Xu, Caiming Xiong, Chris Callison-Burch, Mark Yatskar and Juan Carlos Niebles .....      | 25363 |
| <i>Proactive Hearing Assistants that Isolate Egocentric Conversations</i>                                                                                |       |
| Guilin Hu, Malek Itani, Tuochao Chen and Shyamnath Gollakota .....                                                                                       | 25378 |
| <i>fLSA: Learning Semantic Structures in Document Collections Using Foundation Models</i>                                                                |       |
| Weijia Xu, Nebojsa Jojic and Nicolas Le Roux .....                                                                                                       | 25396 |
| <i>SafeKey: Amplifying Aha-Moment Insights for Safety Reasoning</i>                                                                                      |       |
| Kaiwen Zhou, Xuandong Zhao, Jayanth Srinivasa, Gaowen Liu, Aosong Feng, Dawn Song and Xin Eric Wang .....                                                | 25408 |
| <i>HypER: Literature-grounded Hypothesis Generation and Distillation with Provenance</i>                                                                 |       |
| Rosni Vasu, Chandrayee Basu, Bhavana Dalvi Mishra, Cristina Sarasua, Peter Clark and Abraham Bernstein .....                                             | 25425 |
| <i>Empowering GraphRAG with Knowledge Filtering and Integration</i>                                                                                      |       |
| Kai Guo, Harry Shomer, Shenglai Zeng, Haoyu Han, Yu Wang and Jiliang Tang .....                                                                          | 25451 |
| <i>Interpretable Mnemonic Generation for Kanji Learning via Expectation-Maximization</i>                                                                 |       |
| Jaewook Lee, Alexander Scarlatos and Andrew Lan .....                                                                                                    | 25466 |
| <i>Refining Attention for Explainable and Noise-Robust Fact-Checking with Transformers</i>                                                               |       |
| Jean-Flavien Bussotti and Paolo Papotti .....                                                                                                            | 25488 |
| <i>Harmful Prompt Laundering: Jailbreaking LLMs with Abductive Styles and Symbolic Encoding</i>                                                          |       |
| Seongho Joo, Hyukhun Koh and Kyomin Jung .....                                                                                                           | 25501 |
| <i>Pathway to Relevance: How Cross-Encoders Implement a Semantic Variant of BM25</i>                                                                     |       |
| Meng Lu, Catherine Chen and Carsten Eickhoff .....                                                                                                       | 25537 |
| <i>Rewarding the Unlikely: Lifting GRPO Beyond Distribution Sharpening</i>                                                                               |       |
| Andre Wang He, Daniel Fried and Sean Welleck .....                                                                                                       | 25560 |
| <i>PhoniTale: Phonologically Grounded Mnemonic Generation for Typologically Distant Language Pairs</i>                                                   |       |
| Sana Kang, Myeongseok Gwon, Su Young Kwon, Jaewook Lee, Andrew Lan, Bhiksha Raj and Rita Singh .....                                                     | 25573 |
| <i>Amulet: Putting Complex Multi-Turn Conversations on the Stand with LLM Juries</i>                                                                     |       |
| Sahana Ramnath, Anurag Mudgil, Brihi Joshi, Skyler Hallinan and Xiang Ren .....                                                                          | 25606 |
| <i>Exploring Chain-of-Thought Reasoning for Steerable Pluralistic Alignment</i>                                                                          |       |
| Yunfan Zhang, Kathleen McKeown and Smaranda Muresan .....                                                                                                | 25648 |

|                                                                                                                                       |       |
|---------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>CMedCalc-Bench: A Fine-Grained Benchmark for Chinese Medical Calculations in LLM</i>                                               |       |
| Yunyan Zhang, Zhihong Zhu and Xian Wu .....                                                                                           | 25662 |
| <i>Evaluating Robustness of Large Audio Language Models to Audio Injection: An Empirical Study</i>                                    |       |
| Guanyu Hou, Jiaming He, Yinhang Zhou, Ji Guo, Yitong Qiao, Rui Zhang and Wenbo Jiang                                                  | 25672 |
| <i>How Far Can LLMs Improve from Experience? Measuring Test-Time Learning Ability in LLMs with Human Comparison</i>                   |       |
| Jiayin Wang, Zhiqiang Guo, Weizhi Ma and Min Zhang .....                                                                              | 25689 |
| <i>Subtle Risks, Critical Failures: A Framework for Diagnosing Physical Safety of LLMs for Embodied Decision Making</i>               |       |
| Yejin Son, Minseo Kim, Sungwoong Kim, Seungju Han, Jian Kim, Dongju Jang, Youngjae Yu and Chan Young Park .....                       | 25704 |
| <i>SwiftKV: Fast Prefill-Optimized Inference with Knowledge-Preserving Model Transformation</i>                                       |       |
| Aurick Qiao, Zhewei Yao, Samyam Rajbhandari and Yuxiong He .....                                                                      | 25746 |
| <i>Co-Eval: Augmenting LLM-based Evaluation with Machine Metrics</i>                                                                  |       |
| Ling-I Wu, Weijie Wu, Minyu Chen, Jianxin Xue and Guoqiang Li .....                                                                   | 25766 |
| <i>Sali4Vid: Saliency-Aware Video Reweighting and Adaptive Caption Retrieval for Dense Video Captioning</i>                           |       |
| MinJu Jeon, Si-Woo Kim, Ye-Chan Kim, HyunGee Kim and Dong-Jin Kim .....                                                               | 25789 |
| <i>Semantic Networks Extracted from Students' Think-Aloud Data are Correlated with Students' Learning Performance</i>                 |       |
| Pingjing Yang, Sullam Jeoung, Jennifer Cromley and Jana Diesner .....                                                                 | 25803 |
| <i>Less is More: The Effectiveness of Compact Typological Language Representations</i>                                                |       |
| York Hay Ng, Phuong Hanh Hoang and En-Shiun Annie Lee .....                                                                           | 25817 |
| <i>Sparse Activation Editing for Reliable Instruction Following in Narratives</i>                                                     |       |
| Runcong Zhao, Chengyu Cao, Qinglin Zhu, Xiucheng Ly, Shun Shao, Lin Gui, Ruifeng Xu and Yulan He .....                                | 25829 |
| <i>Inceptive Transformers: Enhancing Contextual Representations through Multi-Scale Feature Learning Across Domains and Languages</i> |       |
| Asif Shahriar, Rifat Shahriyar and M Saifur Rahman .....                                                                              | 25845 |
| <i>Causal Tree Extraction from Medical Case Reports: A Novel Task for Experts-like Text Comprehension</i>                             |       |
| Sakiko Yahata, Zhen Wan, Fei Cheng, Sadao Kurohashi, Hisahiko Sato and Ryoza Nagai .                                                  | 25861 |
| <i>OWL: Probing Cross-Lingual Recall of Memorized Texts via World Literature</i>                                                      |       |
| Alisha Srivastava, Emir Kaan Korukluoglu, Minh Nhat Le, Duyen Tran, Chau Minh Pham, Marzena Karpinska and Mohit Iyyer .....           | 25880 |
| <i>Enhanced Noun-Noun Compound Interpretation through Textual Enrichment</i>                                                          |       |
| Bingyang Ye, Jingxuan Tu and James Pustejovsky .....                                                                                  | 25908 |
| <i>ICL CIPHERS: Quantifying "Learning" in In-Context Learning via Substitution Ciphers</i>                                            |       |
| Zhouxiang Fang, Aayush Mishra, Muhan Gao, Anqi Liu and Daniel Khashabi .....                                                          | 25924 |
| <i>Corrupted but Not Broken: Understanding and Mitigating the Negative Impacts of Corrupted Data in Visual Instruction Tuning</i>     |       |
| Yunhao Gou, Hansi Yang, Zhili Liu, Kai Chen, Yihan Zeng, Lanqing Hong, Zhenguo Li, Qun Liu, Bo Han, James Kwok and Yu Zhang .....     | 25946 |



|                                                                                                                                   |       |
|-----------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Memory OS of AI Agent</i>                                                                                                      |       |
| Jiazheng Kang, Mingming Ji, Zhe Zhao and Ting Bai .....                                                                           | 25973 |
| <i>Rule Discovery for Natural Language Inference Data Generation Using Out-of-Distribution Detection</i>                          |       |
| Juyoung Han, Hyunsun Hwang and Changki Lee .....                                                                                  | 25983 |
| <i>Jigsaw-Puzzles: From Seeing to Understanding to Reasoning in Vision-Language Models</i>                                        |       |
| Zesen Lyu, Dandan Zhang, Wei Ye, Fangdi Li, Zhihang Jiang and Yao Yang .....                                                      | 26004 |
| <i>Definition Generation for Word Meaning Modeling: Monolingual, Multilingual, and Cross-Lingual Perspectives</i>                 |       |
| Francesco Periti, Roksana Goworek, Haim Dubossarsky and Nina Tahmasebi .....                                                      | 26016 |
| <i>Language Model Based Text-to-Audio Generation: Anti-Causally Aligned Collaborative Residual Transformers</i>                   |       |
| Juncheng Wang, Chao Xu, Cheng Yu, Zhe Hu, Haoyu Xie, Guoqi Yu, Lei Shang and Shujun Wang .....                                    | 26037 |
| <i>HELENE: Hessian Layer-wise Clipping and Gradient Annealing for Accelerating Fine-tuning LLM with Zeroth-order Optimization</i> |       |
| Huaqin Zhao, Jiaxi Li, Yi Pan, Shizhe Liang, Xiaofeng Yang, Fei Dou, Tianming Liu and Jin Lu .....                                | 26056 |
| <i>Zero-shot Multimodal Document Retrieval via Cross-modal Question Generation</i>                                                |       |
| Yejin Choi, Jaewoo Park, Janghan Yoon, Saejin Kim, Jaehyun Jeon and Youngjae Yu .....                                             | 26080 |
| <i>From Parameters to Performance: A Data-Driven Study on LLM Structure and Development</i>                                       |       |
| Suqing Wang, Zuchao Li, Shi Luohe, Bo Du, Hai Zhao, Yun Li and Qianren Wang .....                                                 | 26096 |
| <i>Logical Reasoning with Outcome Reward Models for Test-Time Scaling</i>                                                         |       |
| Ramya Keerthy Thatikonda, Wray Buntine and Ehsan Shareghi .....                                                                   | 26114 |
| <i>Speculating LLMs' Chinese Training Data Pollution from Their Tokens</i>                                                        |       |
| Qingjie Zhang, Di Wang, Haoting Qian, Liu Yan, Tianwei Zhang, Ke Xu, Qi Li, Minlie Huang, Hewu Li and Han Qiu .....               | 26125 |
| <i>NovelHopQA: Diagnosing Multi-Hop Reasoning Failures in Long Narrative Contexts</i>                                             |       |
| Abhay Gupta, Kevin Zhu, Vasu Sharma, Sean O'Brien and Michael Lu .....                                                            | 26146 |
| <i>Weights-Rotated Preference Optimization for Large Language Models</i>                                                          |       |
| Chenxu Yang, Ruipeng Jia, Mingyu Zheng, Naibin Gu, Zheng Lin, Siyuan Chen, Weichong Yin, Hua Wu and Weiping Wang .....            | 26164 |
| <i>The Stepwise Deception: Simulating the Evolution from True News to Fake News with LLM Agents</i>                               |       |
| Yuhan Liu, Zirui Song, Juntian Zhang, Xiaoqing Zhang, Xiuying Chen and Rui Yan .....                                              | 26188 |
| <i>How to inject knowledge efficiently? Knowledge Infusion Scaling Law for Pre-training Large Language Models</i>                 |       |
| Kangtao Lv, Haibin Chen, Yujin Yuan, Langming Liu, Shilei Liu, Yongwei Wang, Wenbo Su and Bo Zheng .....                          | 26205 |
| <i>SMEC: Rethinking Matryoshka Representation Learning for Retrieval Embedding Compression</i>                                    |       |
| Biao Zhang, Lixin Chen, Tong Liu and Bo Zheng .....                                                                               | 26221 |
| <i>Reverse Prompt Engineering: A Zero-Shot, Genetic Algorithm Approach to Language Model Inversion</i>                            |       |
| Hanqing Li and Diego Klabjan .....                                                                                                | 26235 |

|                                                                                                                                               |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>DiMo-GUI: Advancing Test-time Scaling in GUI Grounding via Modality-Aware Visual Reasoning</i>                                             |       |
| Hang Wu, Hongkai Chen, Yujun Cai, Chang Liu, Qingwen Ye, Ming-Hsuan Yang and Yiwei Wang .....                                                 | 26258 |
| <i>SocioBench: Modeling Human Behavior in Sociological Surveys with Large Language Models</i>                                                 |       |
| Jia Wang, Ziyu Zhao, Tingjuntao Ni and Zhongyu Wei .....                                                                                      | 26269 |
| <i>Financial Risk Relation Identification through Dual-view Adaptation</i>                                                                    |       |
| Wei-Ning Chiu, Yu-Hsiang Wang, Andy Hsiao, Yu-Shiang Huang and Chuan-Ju Wang ...                                                              | 26302 |
| <i>CopySpec: Accelerating LLMs with Speculative Copy-and-Paste</i>                                                                            |       |
| Razvan-Gabriel Dumitru, Minglai Yang, Vikas Yadav and Mihai Surdeanu .....                                                                    | 26313 |
| <i>GRASP: Replace Redundant Layers with Adaptive Singular Parameters for Efficient Model Compression</i>                                      |       |
| Kainan Liu, Yong Zhang, Ning Cheng, Zhitao Li, Shaojun Wang and Jing Xiao .....                                                               | 26345 |
| <i>GraphAgent: Agentic Graph Language Assistant</i>                                                                                           |       |
| Yuhao Yang, Jiabin Tang, Lianghao Xia, Xingchen Zou, Yuxuan Liang and Chao Huang ..                                                           | 26361 |
| <i>DDO: Dual-Decision Optimization for LLM-Based Medical Consultation via Multi-Agent Collaboration</i>                                       |       |
| Zhihao Jia, Mingyi Jia, Junwen Duan and Jianxin Wang .....                                                                                    | 26381 |
| <i>FedMABench: Benchmarking Mobile GUI Agents on Decentralized Heterogeneous User Data</i>                                                    |       |
| WenHao Wang, Zijie Yu, Rui Ye, Jianqing Zhang, Guangyi Liu, Liang Liu, Siheng Chen and Yanfeng Wang .....                                     | 26399 |
| <i>VLA-Mark: A cross modal watermark for large vision-language alignment models</i>                                                           |       |
| Shuliang Liu, Zheng Qi, Jesse Jiayi Xu, Yibo Yan, Junyan Zhang, He Geng, Aiwei Liu, Peijie Jiang, Jia Liu, Yik-Cheung Tam and Xuming Hu ..... | 26421 |
| <i>Sentence Smith: Controllable Edits for Evaluating Text Embeddings</i>                                                                      |       |
| Hongji Li, Andrianos Michail, Reto Gubelmann, Simon Clematide and Juri Opitz .....                                                            | 26440 |
| <i>ReasonMed: A 370K Multi-Agent Generated Dataset for Advancing Medical Reasoning</i>                                                        |       |
| Yu Sun, Xingyu Qian, Weiwen Xu, Hao Zhang, Chenghao Xiao, Long Li, Deli Zhao, Wenbing Huang, Tingyang Xu, Qifeng Bai and Yu Rong .....        | 26458 |
| <i>Decoding Dense Embeddings: Sparse Autoencoders for Interpreting and Discretizing Dense Retrieval</i>                                       |       |
| Seongwan Park, Taeklim Kim and Youngjoong Ko .....                                                                                            | 26480 |
| <i>UICOMPASS: UI Map Guided Mobile Task Automation via Adaptive Action Generation</i>                                                         |       |
| Yuanzhang Lin, Zhe Zhang, He Rui, Qingao Dong, Mingyi Zhou, Jing Zhang, Xiang Gao and Hailong Sun .....                                       | 26498 |
| <i>Leaky Thoughts: Large Reasoning Models Are Not Private Thinkers</i>                                                                        |       |
| Tommaso Green, Martin Gubri, Haritz Puerto, Sangdoo Yun and Seong Joon Oh .....                                                               | 26519 |
| <i>Model Unlearning via Sparse Autoencoder Subspace Guided Projections</i>                                                                    |       |
| Xu Wang, Zihao Li, Benyou Wang, Yan Hu and Difan Zou .....                                                                                    | 26542 |
| <i>ConvSearch-RI: Enhancing Query Reformulation for Conversational Search with Reasoning via Reinforcement Learning</i>                       |       |
| Changtai Zhu, Siyin Wang, Ruijun Feng, Kai Song and Xipeng Qiu .....                                                                          | 26559 |



|                                                                                                                              |       |
|------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>How to Make Large Language Models Generate 100% Valid Molecules?</i>                                                      |       |
| Wen Tao, Jing Tang, Alvin Chan, Bryan Hooi, Baolong Bi, Nanyun Peng, Yuansheng Liu and Yiwei Wang .....                      | 26577 |
| <i>Exploring Quality and Diversity in Synthetic Data Generation for Argument Mining</i>                                      |       |
| Jianzhu Bao, Yuqi Huang, Yang Sun, Wenya Wang, Yice Zhang, Bojun Jin and Ruifeng Xu                                          | 26593 |
| <i>Dynamic Jointly Batch Selection for Data Efficient Machine Translation Fine-Tuning</i>                                    |       |
| Mohammad Amin Ghanizadeh and Mohammad Javad Dousti .....                                                                     | 26617 |
| <i>3MDBench: Medical Multimodal Multi-agent Dialogue Benchmark</i>                                                           |       |
| Ivan Sviridov, Amina Miftakhova, Tereshchenko Artemiy Vladimirovich, Galina Zubkova, Pavel Blinov and Andrey Savchenko ..... | 26626 |
| <i>OpenTuringBench: An Open-Model-based Benchmark and Framework for Machine-Generated Text Detection and Attribution</i>     |       |
| Lucio La Cava and Andrea Tagarelli .....                                                                                     | 26667 |
| <i>CRITICTOOL: Evaluating Self-Critique Capabilities of Large Language Models in Tool-Calling Error Scenarios</i>            |       |
| Shiting Huang, Zhen Fang, Zehui Chen, Siyu Yuan, Junjie Ye, Yu Zeng, Lin Chen, Qi Mao and Feng Zhao .....                    | 26684 |
| <i>Pre-trained Language Models Learn Remarkably Accurate Representations of Numbers</i>                                      |       |
| Marek Kadlčík, Michal Štefánik, Timothee Mickus, Josef Kuchař and Michal Spiegel ...                                         | 26694 |
| <i>Enhancing Large Vision-Language Models with Ultra-Detailed Image Caption Generation</i>                                   |       |
| Yu Zeng, Yukun Qi, Yiming Zhao, Xikun Bao, Lin Chen, Zehui Chen, Shiting Huang, Jie Zhao and Feng Zhao .....                 | 26704 |
| <i>Translate Smart, not Hard: Cascaded Translation Systems with Quality-Aware Deferral</i>                                   |       |
| António Farinhas, Nuno M Guerreiro, Sweta Agrawal, Ricardo Rei and Andre Martins ...                                         | 26731 |
| <i>iVISPAR — An Interactive Visual-Spatial Reasoning Benchmark for VLMs</i>                                                  |       |
| Julius Mayer, Mohamad Ballout, Serwan Jassim, Farbod Nosrat Nezami and Elia Bruni ...                                        | 26746 |
| <i>Are LLMs Better than Reported? Detecting Label Errors and Mitigating Their Effect on Model Performance</i>                |       |
| Omer Nahum, Nitay Calderon, Orgad Keller, Idan Szpektor and Roi Reichart .....                                               | 26771 |
| <i>Detecting Legal Citations in United Kingdom Court Judgments</i>                                                           |       |
| Holli Sargeant, Andreas Östling and Måns Magnusson .....                                                                     | 26799 |
| <i>Large Language Models Badly Generalize across Option Length, Problem Types, and Irrelevant Noun Replacements</i>          |       |
| Guangxiang Zhao, Saier Hu, Xiaoqi Jian, Wu Jinzhu, Yuhan Wu, Lin Sun and Xiangzheng Zhang                                    | 26826 |
| <i>Studying the Role of Input-Neighbor Overlap in Retrieval-Augmented Language Models Training Efficiency</i>                |       |
| Ehsan Doostmohammadi and Marco Kuhlmann .....                                                                                | 26836 |
| <i>Principled Personas: Defining and Measuring the Intended Effects of Persona Prompting on Task Performance</i>             |       |
| Pedro Henrique Luz de Araujo, Paul Röttger, Dirk Hovy and Benjamin Roth .....                                                | 26846 |

|                                                                                                                                                    |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>HydraOpt: Navigating the Efficiency-Performance Trade-off of Adapter Merging</i>                                                                |       |
| Taha Ceritli, Ondrej Bohdal, Mete Ozay, Jijoong Moon, Kyenghun Lee, Hyeonmok Ko and Um-<br>berto Michieli .....                                    | 26876 |
| <i>Parrot: A Training Pipeline Enhances Both Program CoT and Natural Language CoT for Reasoning</i>                                                |       |
| Senjie Jin, Lu Chen, Zhiheng Xi, Yuhui Wang, Sirui Song, Yuhao Zhou, Xinbo Zhang, Peng Sun,<br>Hong Lu, Tao Gui, Qi Zhang and Xuanjing Huang ..... | 26899 |
| <i>Spec-VLA: Speculative Decoding for Vision-Language-Action Models with Relaxed Acceptance</i>                                                    |       |
| Songsheng Wang, Rucheng Yu, Zhihang Yuan, Chao Yu, Feng Gao, Yu Wang and Derek F. Wong                                                             | 26917 |
| <i>Leveraging Text-to-Text Transformers as Classifier Chain for Few-Shot Multi-Label Classification</i>                                            |       |
| Quang Anh Nguyen, Nadi Tomeh, Mustapha Lebbah, Thierry Charnois and Hanane Azzag                                                                   | 26930 |
| <i>M-Wanda: Improving One-Shot Pruning for Multilingual LLMs</i>                                                                                   |       |
| Rochelle Choenni and Ivan Titov .....                                                                                                              | 26940 |
| <i>Beyond Hate Speech: NLP's Challenges and Opportunities in Uncovering Dehumanizing Language</i>                                                  |       |
| Hamidreza Saffari, Mohammadamin Shafiei, Hezhao Zhang, Lasana T. Harris and Nafise Sadat<br>Moosavi .....                                          | 26954 |
| <i>Conflict-Aware Soft Prompting for Retrieval-Augmented Generation</i>                                                                            |       |
| Eunseong Choi, June Park, Hyeri Lee and Jongwuk Lee .....                                                                                          | 26970 |
| <i>R-CHAR: A Metacognition-Driven Framework for Role-Playing in Large Language Models</i>                                                          |       |
| Haiming Qin, Jiwei Zhang, Wei Zhang, KeZhong Lu, Mingyang Zhou, Hao Liao and Rui Mao                                                               | 26985 |
| <i>Annotating Training Data for Conditional Semantic Textual Similarity Measurement using Large Lan-<br/>guage Models</i>                          |       |
| Gaifan Zhang, Yi Zhou and Danushka Bollegala .....                                                                                                 | 27004 |
| <i>When Words Smile: Generating Diverse Emotional Facial Expressions from Text</i>                                                                 |       |
| Haidong Xu, Meishan Zhang, Hao Ju, Zhedong Zheng, Erik Cambria, Min Zhang and Hao Fei                                                              | 27017 |
| <i>Improving Online Job Advertisement Analysis via Compositional Entity Extraction</i>                                                             |       |
| Kai Krüger, Johanna Binnewitt, Kathrin Ehmann, Stefan Winnige and Alan Akbik .....                                                                 | 27036 |
| <i>Correlation-Aware Example Selection for In-Context Learning with Nonsymmetric Determinantal Point<br/>Processes</i>                             |       |
| Qiunan Du, Zhiliang Tian, Zhen Huang, Kailun Bian, Tianlun Liu, Zhaoning Zhang, Xinwang<br>Liu, Feng Liu and Dongsheng Li .....                    | 27055 |
| <i>Leveraging Cognitive Complexity of Texts for Contextualization in Dense Retrieval</i>                                                           |       |
| Effrosyni Sokli, Georgios Peikos, Pranav Kasela and Gabriella Pasi .....                                                                           | 27072 |
| <i>Beyond Online Sampling: Bridging Offline-to-Online Alignment via Dynamic Data Transformation for<br/>LLMs</i>                                   |       |
| Zhang Zhang, Guhao Feng, Jian Guan, Di He and Wei Wu .....                                                                                         | 27086 |
| <i>CAVE : Detecting and Explaining Commonsense Anomalies in Visual Environments</i>                                                                |       |
| Rishika Bhagwatkar, Syrielle Montariol, Angelika Romanou, Beatriz Borges, Irina Rish and An-<br>toine Bosselut .....                               | 27099 |

|                                                                                                                                                                            |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Enhancing LLM Language Adaption through Cross-lingual In-Context Pre-training</i>                                                                                       |       |
| Linjuan Wu, Hao-Ran Wei, Huan Lin, Tianhao Li, Baosong Yang, Fei Huang and Weiming Lu                                                                                      |       |
|                                                                                                                                                                            | 27141 |
| <i>SemVink: Advancing VLMs' Semantic Understanding of Optical Illusions via Visual Global Thinking</i>                                                                     |       |
| Sifan Li, Yujun Cai and Yiwei Wang                                                                                                                                         | 27156 |
| <i>Order Doesn't Matter, But Reasoning Does: Training LLMs with Order-Centric Augmentation</i>                                                                             |       |
| Qianxi He, Qianyu He, Jiaqing Liang, Weikang Zhou, Zeye Sun, Fei Yu and Yanghua Xiao                                                                                       | 27167 |
| <i>Type-Less yet Type-Aware Inductive Link Prediction with Pretrained Language Models</i>                                                                                  |       |
| Alessandro De Bellis, Salvatore Bui, Giovanni Servedio, Vito Walter Anelli, Tommaso Di Noia and Eugenio Di Sciascio                                                        | 27182 |
| <i>Extracting Linguistic Information from Large Language Models: Syntactic Relations and Derivational Knowledge</i>                                                        |       |
| Tsedeniya Kinfe Temesgen, Marion Di Marco and Alexander Fraser                                                                                                             | 27199 |
| <i>Beyond Correctness: Confidence-Aware Reward Modeling for Enhancing Large Language Model Reasoning</i>                                                                   |       |
| Qianxi He, Qingyu Ren, Shanzhe Lei, Xuhong Wang and Yingchun Wang                                                                                                          | 27216 |
| <i>TrojanStego: Your Language Model Can Secretly Be A Steganographic Privacy Leaking Agent</i>                                                                             |       |
| Dominik Meier, Jan Philip Wahle, Paul Röttger, Terry Ruas and Bela Gipp                                                                                                    | 27233 |
| <i>Frequency &amp; Compositionality in Emergent Communication</i>                                                                                                          |       |
| Jean-Baptiste Sevestre and Emmanuel Dupoux                                                                                                                                 | 27251 |
| <i>Summarizing Speech: A Comprehensive Survey</i>                                                                                                                          |       |
| Fabian Retkowsky, Maike Züfle, Andreas Sudmann, Dinah Pfau, Shinji Watanabe, Jan Niehues and Alexander Waibel                                                              | 27264 |
| <i>CogDual: Enhancing Dual Cognition of LLMs via Reinforcement Learning with Implicit Rule-Based Rewards</i>                                                               |       |
| Cheng Liu, Yifei Lu, Fanghua Ye, Jian Li, Xingyu Chen, Feiliang Ren, Zhaopeng Tu and Xiaolong Li                                                                           | 27296 |
| <i>Assay2Mol: Large Language Model-based Drug Design Using BioAssay Context</i>                                                                                            |       |
| Yifan Deng, Spencer S Ericksen and Anthony Gitter                                                                                                                          | 27326 |
| <i>Frame First, Then Extract: A Frame-Semantic Reasoning Pipeline for Zero-Shot Relation Triplet Extraction</i>                                                            |       |
| Zehan Li, Fu Zhang, Wenqing Zhang, Jiawei Li, Zhou Li, Jingwei Cheng and Tianyue Peng                                                                                      | 27352 |
| <i>MrGuard: A Multilingual Reasoning Guardrail for Universal LLM Safety</i>                                                                                                |       |
| Yahan Yang, Soham Dan, Shuo Li, Dan Roth and Insup Lee                                                                                                                     | 27366 |
| <i>TALON: A Multi-Agent Framework for Long-Table Exploration and Question Answering</i>                                                                                    |       |
| Ruochun Jin, Xiyue Wang, Dong Wang, Haoqi Zheng, Yunpeng Qi, Silin Yang and Meng Zhang                                                                                     | 27386 |
| <i>You Are What You Train: Effects of Data Composition on Training Context-aware Machine Translation Models</i>                                                            |       |
| Pawel Maka, Yusuf Can Semerci, Jan Scholtes and Gerasimos Spanakis                                                                                                         | 27403 |
| <i>Improving Neutral Point-of-View Generation with Data- and Parameter-Efficient RL</i>                                                                                    |       |
| Jessica Hoffmann, Christiane Ahlheim, Zac Yu, Aria Walfrand, Jarvis Jin, Marie Tano, Ahmad Beirami, Erin MacMurray van Liemt, Nithum Thain, Hakim Sidahmed and Lucas Dixon | 27427 |

|                                                                                                                                                                                                                                                                                                                                                      |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Randomized Smoothing Meets Vision-Language Models</i>                                                                                                                                                                                                                                                                                             |       |
| Emmanouil Seferis, Changshun Wu, Stefanos Kollias, Saddek Bensalem and Chih-Hong Cheng                                                                                                                                                                                                                                                               |       |
|                                                                                                                                                                                                                                                                                                                                                      | 27457 |
| <i>PIIvot: A Lightweight NLP Anonymization Framework for Question-Anchored Tutoring Dialogues</i>                                                                                                                                                                                                                                                    |       |
| Matthew Zent, Digory Smith and Simon Woodhead                                                                                                                                                                                                                                                                                                        | 27468 |
| <i>Trustworthy Medical Question Answering: An Evaluation-Centric Survey</i>                                                                                                                                                                                                                                                                          |       |
| Yinuo Wang, Baiyang Wang, Robert Mercer, Frank Rudzicz, Sudipta Singha Roy, Pengjie Ren, Zhumin Chen and Xindi Wang                                                                                                                                                                                                                                  | 27478 |
| <i>Unpacking Let Alone: Human-Scale Models Generalize to a Rare Construction in Form but not Meaning</i>                                                                                                                                                                                                                                             |       |
| Wesley Scivetti, Tatsuya Aoyama, Ethan Wilcox and Nathan Schneider                                                                                                                                                                                                                                                                                   | 27492 |
| <i>BOUQuET : dataset, Benchmark and Open initiative for Universal Quality Evaluation in Translation</i>                                                                                                                                                                                                                                              |       |
| Pierre Andrews, Mikel Artetxe, Mariano Coria Meglioli, Marta R. Costa-jussà, Joe Chuang, David Dale, Mark Duppenthaler, Nathaniel Paul Ekberg, Cynthia Gao, Daniel Edward Licht, Jean Maillard, Alexandre Mourachko, Christophe Ropers, Safiyyah Saleem, Eduardo Sánchez, Ioannis Tsiamas, Arina Turkatenco, Albert Ventayol-Boada and Shireen Yates | 27504 |
| <i>HealthCards: Exploring Text-to-Image Generation as Visual Aids for Healthcare Knowledge Democratizing and Education</i>                                                                                                                                                                                                                           |       |
| Qian Wu, Zheyao Gao, Longfei Gou, Yifan Hou, Ann Sin Nga Lau and Qi Dou                                                                                                                                                                                                                                                                              | 27525 |
| <i>When Life Gives You Samples: The Benefits of Scaling up Inference Compute for Multilingual LLMs</i>                                                                                                                                                                                                                                               |       |
| Ammar Khairi, Daniel D'souza, Ye Shen, Julia Kreutzer and Sara Hooker                                                                                                                                                                                                                                                                                | 27548 |
| <i>Creativity in LLM-based Multi-Agent Systems: A Survey</i>                                                                                                                                                                                                                                                                                         |       |
| Yi-Cheng Lin, Kang-Chieh Chen, Zhe-Yan Li, Tzu-Heng Wu, Tzu-Hsuan Wu, Kuan-Yu Chen, Hung-yi Lee and Yun-Nung Chen                                                                                                                                                                                                                                    | 27573 |
| <i>Context and POS in Action: A Comparative Study of Chinese Homonym Disambiguation in Human and Language Models</i>                                                                                                                                                                                                                                 |       |
| Xie Chenwei, Matthew King-Hang Ma, Wenbo Wang and William Shiyuan Wang                                                                                                                                                                                                                                                                               | 27597 |
| <i>Attacking Misinformation Detection Using Adversarial Examples Generated by Language Models</i>                                                                                                                                                                                                                                                    |       |
| Piotr Przybyła, Euan McGill and Horacio Saggion                                                                                                                                                                                                                                                                                                      | 27615 |
| <i>Leveraging Loanword Constraints for Improving Machine Translation in a Low-Resource Multilingual Context</i>                                                                                                                                                                                                                                      |       |
| Felermirino D. M. A. Ali, Henrique Lopes Cardoso and Rui Sousa-Silva                                                                                                                                                                                                                                                                                 | 27632 |
| <i>Linguistic Neuron Overlap Patterns to Facilitate Cross-lingual Transfer on Low-resource Languages</i>                                                                                                                                                                                                                                             |       |
| Yuemei Xu, Kexin Xu, Jian Zhou, Ling Hu and Lin Gui                                                                                                                                                                                                                                                                                                  | 27647 |
| <i>Scaling Low-Resource MT via Synthetic Data Generation with LLMs</i>                                                                                                                                                                                                                                                                               |       |
| Ona de Gibert, Joseph Attieh, Teemu Vahtola, Mikko Aulamo, Zihao Li, Raúl Vázquez, Tiancheng Hu and Jörg Tiedemann                                                                                                                                                                                                                                   | 27663 |
| <i>Tailoring Table Retrieval from a Field-aware Hybrid Matching Perspective</i>                                                                                                                                                                                                                                                                      |       |
| Da Li, Keping Bi, Jiafeng Guo and Xueqi Cheng                                                                                                                                                                                                                                                                                                        | 27682 |
| <i>Randomly Removing 50% of Dimensions in Text Embeddings has Minimal Impact on Retrieval and Classification Tasks</i>                                                                                                                                                                                                                               |       |
| Sotaro Takeshita, Yurina Takeshita, Daniel Ruffinelli and Simone Paolo Ponzetto                                                                                                                                                                                                                                                                      | 27694 |

|                                                                                                                                                                                                                                                                                                           |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Morables: A Benchmark for Assessing Abstract Moral Reasoning in LLMs with Fables</i>                                                                                                                                                                                                                   |       |
| Matteo Marcuzzo, Alessandro Zangari, Andrea Albarelli, Jose Camacho-Collados and Mohammad Taher Pilehvar .....                                                                                                                                                                                            | 27716 |
| <i>MessIRve: A Large-Scale Spanish Information Retrieval Dataset</i>                                                                                                                                                                                                                                      |       |
| Francisco Valentini, Viviana Cotik, Damián Furman, Ivan Bercovich, Edgar Altszyler and Juan Manuel Pérez .....                                                                                                                                                                                            | 27741 |
| <i>AFRIDOC-MT: Document-level MT Corpus for African Languages</i>                                                                                                                                                                                                                                         |       |
| Jesujoba Oluwadara Alabi, Israel Abebe Azime, Miaoran Zhang, Cristina España-Bonet, Rachel Bawden, Dawei Zhu, David Ifeoluwa Adelani, Clement Oyeleke Odoje, Idris Akinade, Iffat Maab, Davis David, Shamsuddeen Hassan Muhammad, Neo Putini, David O. Ademuyiwa, Andrew Caines and Dietrich Klakow ..... | 27759 |
| <i>Charting the Landscape of African NLP: Mapping Progress and Shaping the Road Ahead</i>                                                                                                                                                                                                                 |       |
| Jesujoba Oluwadara Alabi, Michael A. Hedderich, David Ifeoluwa Adelani and Dietrich Klakow .....                                                                                                                                                                                                          | 27796 |
| <i>GLIMPSE: Do Large Vision-Language Models Truly Think With Videos or Just Glimpse at Them?</i>                                                                                                                                                                                                          |       |
| Yiyang Zhou, Linjie Li, Shi Qiu, Zhengyuan Yang, Yuyang Zhao, Siwei Han, Yangfan He, Kangqi Li, Haonian Ji, Zihao Zhao, Haibo Tong, Lijuan Wang and Huaxiu Yao .....                                                                                                                                      | 27831 |
| <i>Social Bias in Multilingual Language Models: A Survey</i>                                                                                                                                                                                                                                              |       |
| Lance Calvin Lim Gamboa, Yue Feng and Mark G. Lee .....                                                                                                                                                                                                                                                   | 27846 |
| <i>BYOKG-RAG: Multi-Strategy Graph Retrieval for Knowledge Graph Question Answering</i>                                                                                                                                                                                                                   |       |
| Costas Mavromatis, Soji Adeshina, Vassilis N. Ioannidis, Zhen Han, Qi Zhu, Ian Robinson, Bryan Thompson, Huzefa Rangwala and George Karypis .....                                                                                                                                                         | 27870 |
| <i>Synth-SBDH: A Synthetic Dataset of Social and Behavioral Determinants of Health for Clinical Text</i>                                                                                                                                                                                                  |       |
| Avijit Mitra, Zhichao Yang, Emily Druhl, Raelene Goodwin and Hong yu .....                                                                                                                                                                                                                                | 27888 |
| <i>Pun Unintended: LLMs and the Illusion of Humor Understanding</i>                                                                                                                                                                                                                                       |       |
| Alessandro Zangari, Matteo Marcuzzo, Andrea Albarelli, Mohammad Taher Pilehvar and Jose Camacho-Collados .....                                                                                                                                                                                            | 27925 |
| <i>RACCoon: Versatile Instructional Video Editing with Auto-Generated Narratives</i>                                                                                                                                                                                                                      |       |
| Jaehong Yoon, Shoubin Yu and Mohit Bansal .....                                                                                                                                                                                                                                                           | 27961 |
| <i>Pre-trained Models Perform the Best When Token Distributions Follow Zipf's Law</i>                                                                                                                                                                                                                     |       |
| Yanjin He, Qingkai Zeng and Meng Jiang .....                                                                                                                                                                                                                                                              | 27998 |
| <i>Do RAG Systems Really Suffer From Positional Bias?</i>                                                                                                                                                                                                                                                 |       |
| Florin Cuconasu, Simone Filice, Guy Horowitz, Yoelle Maarek and Fabrizio Silvestri .....                                                                                                                                                                                                                  | 28011 |
| <i>Aspect-Oriented Summarization for Psychiatric Short-Term Readmission Prediction</i>                                                                                                                                                                                                                    |       |
| WonJin Yoon, Boyu Ren, Spencer Thomas, Chanhwi Kim, Guergana K Savova, Mei-Hua Hall and Timothy A. Miller .....                                                                                                                                                                                           | 28026 |
| <i>Adapting Bias Evaluation to Domain Contexts using Generative Models</i>                                                                                                                                                                                                                                |       |
| Tamara Quiroga, Felipe Bravo-Marquez and Valentin Barriere .....                                                                                                                                                                                                                                          | 28044 |
| <i>Emergent morpho-phonological representations in self-supervised speech models</i>                                                                                                                                                                                                                      |       |
| Jon Gauthier, Canaan Breiss, Matthew K Leonard and Edward F. Chang .....                                                                                                                                                                                                                                  | 28056 |
| <i>Multilingual Language Model Pretraining using Machine-translated Data</i>                                                                                                                                                                                                                              |       |
| Jiayi Wang, Yao Lu, Maurice Weber, Max Ryabinin, David Ifeoluwa Adelani, Yihong Chen, Raphael Tang and Pontus Stenetorp .....                                                                                                                                                                             | 28076 |

|                                                                                                                                     |       |
|-------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>IntentionFrame: A Semi-Structured, Multi-Aspect Framework for Fine-Grained Conversational Intention Understanding</i>            |       |
| Jinggui Liang, Dung Vo and Lizi Liao .....                                                                                          | 28097 |
| <i>Video-RTS: Rethinking Reinforcement Learning and Test-Time Scaling for Efficient and Enhanced Video Reasoning</i>                |       |
| Ziyang Wang, Jaehong Yoon, Shoubin Yu, Md Mohaiminul Islam, Gedas Bertasius and Mohit Bansal .....                                  | 28115 |
| <i>Efficient Compositional Multi-tasking for On-device Large Language Models</i>                                                    |       |
| Ondrej Bohdal, Mete Ozay, Jijoong Moon, Kyenghun Lee, Hyeonmok Ko and Umberto Michieli                                              | 28130 |
| <i>Improving Large Language Model Safety with Contrastive Representation Learning</i>                                               |       |
| Samuel Simko, Mrinmaya Sachan, Bernhard Schölkopf and Zhijing Jin .....                                                             | 28155 |
| <i>Leveraging What's Overfixed: Post-Correction via LLM Grammatical Error Overcorrection</i>                                        |       |
| Taehee Park, Heejin Do and Gary Lee .....                                                                                           | 28184 |
| <i>Scaling Up Temporal Domain Generalization via Temporal Experts Averaging</i>                                                     |       |
| Aoming Liu, Kevin Miller, Venkatesh Saligrama, Kate Saenko, Boqing Gong, Ser-Nam Lim and Bryan A. Plummer .....                     | 28197 |
| <i>LinguaLens: Towards Interpreting Linguistic Mechanisms of Large Language Models via Sparse Auto-Encoder</i>                      |       |
| Yi Jing, Zijun Yao, Hongzhu Guo, Lingxu Ran, Xiaozhi Wang, Lei Hou and Juanzi Li ....                                               | 28221 |
| <i>The Strawberry Problem: Emergence of Character-level Understanding in Tokenized Language Models</i>                              |       |
| Adrian Cosma, Stefan Ruseti, Emilian Radoi and Mihai Dascalu .....                                                                  | 28241 |
| <i>Improving the Quality of Web-mined Parallel Corpora of Low-Resource Languages using Debiasing Heuristics</i>                     |       |
| Surangika Ranathunga, Aloka Fernando, Menan Velayuthan, Charitha Rathnayaka and Nisansa de Silva .....                              | 28253 |
| <i>Weaver: Interweaving SQL and LLM for Table Reasoning</i>                                                                         |       |
| Rohit Khoja, Devanshu Gupta, Yanjie Fu, Dan Roth and Vivek Gupta .....                                                              | 28271 |
| <i>ECO Decoding: Entropy-Based Control for Controllability and Fluency in Controllable Dialogue Generation</i>                      |       |
| Seungmin Shin, Dooyoung Kim and Youngjoong Ko .....                                                                                 | 28298 |
| <i>Investigating the interaction of linguistic and mathematical reasoning in language models using multi-lingual number puzzles</i> |       |
| Antara Raaghavi Bhattacharya, Isabel Papadimitriou, Kathryn Davidson and David Alvarez-Melis                                        | 28311 |
| <i>Unsupervised Concept Vector Extraction for Bias Control in LLMs</i>                                                              |       |
| Hannah Cyberek, Yangfeng Ji and David Evans .....                                                                                   | 28322 |
| <i>Seeing the Same Story Differently: Framing-Divergent Event Coreference for Computational Framing Analysis</i>                    |       |
| Jin Zhao, Xinrui Hu and Nianwen Xue .....                                                                                           | 28345 |
| <i>LLMs are Better Than You Think: Label-Guided In-Context Learning for Named Entity Recognition</i>                                |       |
| Fan Bai, Hamid Hassanzadeh, Ardavan Saeedi and Mark Dredze .....                                                                    | 28361 |



|                                                                                                                                                                     |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>COUNTDOWN: Contextually Sparse Activation Filtering Out Unnecessary Weights in Down Projection</i>                                                               |       |
| Jaewon Cheon and Pilsung Kang .....                                                                                                                                 | 28382 |
| <i>SimpleDoc: Multi-Modal Document Understanding with Dual-Cue Page Retrieval and Iterative Refinement</i>                                                          |       |
| Chelsi Jain, Yiran Wu, Yifan Zeng, Jiale Liu, Shengyu Dai, Zhenwen Shao, Qingyun Wu and Huazheng Wang .....                                                         | 28399 |
| <i>VLP: Vision-Language Preference Learning for Embodied Manipulation</i>                                                                                           |       |
| Runze Liu, Chenjia Bai, Jiafei Lyu, Shengjie Sun, Yali Du and Xiu Li .....                                                                                          | 28417 |
| <i>QG-CoC: Question-Guided Chain-of-Captions for Large Multimodal Models</i>                                                                                        |       |
| Kuei-Chun Kao, Hsu Tzu-Yin, Yunqi Hong, Ruochen Wang and Cho-Jui Hsieh .....                                                                                        | 28434 |
| <i>EGOILLUSION: Benchmarking Hallucinations in Egocentric Video Understanding</i>                                                                                   |       |
| Ashish Seth, Utkarsh Tyagi, Ramaneswaran Selvakumar, Nishit Anand, Sonal Kumar, Sreyan Ghosh, Ramani Duraiswami, Chirag Agarwal and Dinesh Manocha .....            | 28450 |
| <i>MULTIVOX: A Benchmark for Evaluating Voice Assistants for Multimodal Interactions</i>                                                                            |       |
| Ramaneswaran Selvakumar, Ashish Seth, Nishit Anand, Utkarsh Tyagi, Sonal Kumar, Sreyan Ghosh and Dinesh Manocha .....                                               | 28470 |
| <i>Do All Autoregressive Transformers Remember Facts the Same Way? A Cross-Architecture Analysis of Recall Mechanisms</i>                                           |       |
| Minyeong Choe, Haehyun Cho, Changho Seo and Hyunil Kim .....                                                                                                        | 28483 |
| <i>Probing Narrative Morals: A New Character-Focused MFT Framework for Use with Large Language Models</i>                                                           |       |
| Luca Mitran, Sophie Wu and Andrew Piper .....                                                                                                                       | 28503 |
| <i>Probing and Boosting Large Language Models Capabilities via Attention Heads</i>                                                                                  |       |
| Dezhi Zhao, Xin Liu, Xiaocheng Feng, Hui Wang and Bing Qin .....                                                                                                    | 28519 |
| <i>A Survey of Link Prediction in N-ary Knowledge Graphs</i>                                                                                                        |       |
| Jiyao Wei, Saiping Guan, Da Li, Zhongni Hou, Miao Su, Yucan Guo, Xiaolong Jin, Jiafeng Guo and Xueqi Cheng .....                                                    | 28534 |
| <i>Multi-Frequency Contrastive Decoding: Alleviating Hallucinations for Large Vision-Language Models</i>                                                            |       |
| Bingqian Liu, Fu Zhang, Guoqing Chen and Jingwei Cheng .....                                                                                                        | 28557 |
| <i>ORPP: Self-Optimizing Role-playing Prompts to Enhance Language Model Capabilities</i>                                                                            |       |
| Yifan Duan, Yihong Tang, Kehai Chen, Liqiang Nie and Min Zhang .....                                                                                                | 28574 |
| <i>BrailleLLM: Braille Instruction Tuning with Large Language Models for Braille Domain Tasks</i>                                                                   |       |
| Tianyuan Huang, Zepeng Zhu, Hangdi Xing, Zirui Shao, Zhi Yu, Chaoxiong Yang, Jiaxian He, Xiaozhong Liu and Jiajun Bu .....                                          | 28590 |
| <i>MaviS: A Multimodal Conversational Assistant For Avian Species</i>                                                                                               |       |
| Yevheniia Kryklyvets, Mohammed Irfan Kurpath, Sahal Shaji Mullappilly, Jinxing Zhou, Fahad Shahbaz Khan, Rao Muhammad Anwer, Salman Khan and Hisham Cholakkal ..... | 28602 |
| <i>Refining Text Generation for Realistic Conversational Recommendation via Direct Preference Optimization</i>                                                      |       |
| Manato Tajiri and Michimasa Inaba .....                                                                                                                             | 28629 |
| <i>Large Language Models Threaten Language's Epistemic and Communicative Foundations</i>                                                                            |       |
| Shashank Srivastava .....                                                                                                                                           | 28651 |

|                                                                                                                                       |       |
|---------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Detecting Knowledge Boundary of Vision Large Language Models by Sampling-Based Inference</i>                                       |       |
| Zhuo Chen, Xinyu Wang, Yong Jiang, Zhen Zhang, Xinyu Geng, Pengjun Xie, Fei Huang and Kewei Tu .....                                  | 28666 |
| <i>Multi-view-guided Passage Reranking with Large Language Models</i>                                                                 |       |
| Jeongwoo Na, Jun Kwon, Eunseong Choi and Jongwuk Lee .....                                                                            | 28682 |
| <i>Disentangling Subjectivity and Uncertainty for Hate Speech Annotation and Modeling using Gaze</i>                                  |       |
| Özge Alacam, Sanne Hoeken, Andreas Säuberli, Hannes Gröner, Diego Frassinelli, Sina Zarrieß and Barbara Plank .....                   | 28696 |
| <i>VoiceBBQ: Investigating Effect of Content and Acoustics in Social Bias of Spoken Language Model</i>                                |       |
| Junhyuk Choi, Ro-hoon Oh, Jihwan Seol and Bugeun Kim .....                                                                            | 28714 |
| <i>Explaining Differences Between Model Pairs in Natural Language through Sample Learning</i>                                         |       |
| Advaith Malladi, Rakesh R Menon, Yuvraj Jain and Shashank Srivastava .....                                                            | 28726 |
| <i>Compound AI Systems Optimization: A Survey of Methods, Challenges, and Future Directions</i>                                       |       |
| Yu-Ang Lee, Guan-Ting Yi, Mei-Yi Liu, Jui-Chao Lu, Guan-Bo Yang and Yun-Nung Chen                                                     | 28749 |
| <i>A Multi-Level Benchmark for Causal Language Understanding in Social Media Discourse</i>                                            |       |
| Xiaohan Ding, Kaike Ping, Buse Çarık and Eugenia Rho .....                                                                            | 28765 |
| <i>Causal Representation Learning from Multimodal Clinical Records under Non-Random Modality Missingness</i>                          |       |
| Zihan Liang, Ziwen Pan and Ruoxuan Xiong .....                                                                                        | 28780 |
| <i>XLQA: A Benchmark for Locale-Aware Multilingual Open-Domain Question Answering</i>                                                 |       |
| Keonwoo Roh, Yeong-Joon Ju and Seong-Wan Lee .....                                                                                    | 28798 |
| <i>Transformer-Based Temporal Information Extraction and Application: A Review</i>                                                    |       |
| Xin Su, Phillip Howard and Steven Bethard .....                                                                                       | 28811 |
| <i>How to Protect Yourself from 5G Radiation? Investigating LLM Responses to Implicit Misinformation</i>                              |       |
| Ruohao Guo, Wei Xu and Alan Ritter .....                                                                                              | 28831 |
| <i>AmpleHate: Amplifying the Attention for Versatile Implicit Hate Detection</i>                                                      |       |
| Yejin Lee, Joonghyuk Hahn, Hyeseon Ahn and Yo-Sub Han .....                                                                           | 28851 |
| <i>Can Large Language Models Act as Ensembler for Multi-GNNs?</i>                                                                     |       |
| Hanqi Duan, Yao Cheng, Jianxiang Yu, Yao Liu and Xiang Li .....                                                                       | 28864 |
| <i>Agent-to-Agent Theory of Mind: Testing Interlocutor Awareness among Large Language Models</i>                                      |       |
| Younwoo Choi, Changling Li, Yongjin Yang and Zhijing Jin .....                                                                        | 28884 |
| <i>From Charts to Fair Narratives: Uncovering and Mitigating Geo-Economic Biases in Chart-to-Text</i>                                 |       |
| Ridwan Mahbub, Mohammed Saidul Islam, Mir Tafseer Nayeem, Md Tahmid Rahman Laskar, Mizanur Rahman, Shafiq Joty and Enamul Hoque ..... | 28918 |
| <i>Real-time Ad Retrieval via LLM-generative Commercial Intention for Sponsored Search Advertising</i>                                |       |
| Tongtong Liu, Zhaohui Wang, Meiyue Qin, Zenghui Lu, Xudong Chen, Yuekui Yang and Peng Shu .....                                       | 28937 |
| <i>Toward Efficient Sparse Autoencoder-Guided Steering for Improved In-Context Learning in Large Language Models</i>                  |       |
| Ikhyun Cho and Julia Hockenmaier .....                                                                                                | 28950 |



|                                                                                                                                                                                                  |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>CLMTracing: Black-box User-level Watermarking for Code Language Model Tracing</i>                                                                                                             |       |
| Boyue Zhang, Ping He, Tianyu Du, Xuhong Zhang, Lei Yun, Kingsum Chow and Jianwei Yin                                                                                                             | 28963 |
| <i>The Good, the Bad and the Constructive: Automatically Measuring Peer Review's Utility for Authors</i>                                                                                         |       |
| Abdelrahman Sadallah, Tim Baumgärtner, Iryna Gurevych and Ted Briscoe                                                                                                                            | 28980 |
| <i>Evolving Chinese Spelling Correction with Corrector-Verifier Collaboration</i>                                                                                                                |       |
| Linfeng Liu, Hongqiu Wu and Hai Zhao                                                                                                                                                             | 29011 |
| <i>M2Edit: Locate and Edit Multi-Granularity Knowledge in Multimodal Large Language Model</i>                                                                                                    |       |
| Yang Zhou, Pengfei Cao, Yubo Chen, Qingbin Liu, Dianbo Sui, Xi Chen, Kang Liu and Jun Zhao                                                                                                       | 29018 |
| <i>Do LLMs Behave as Claimed? Investigating How LLMs Follow Their Own Claims using Counterfactual Questions</i>                                                                                  |       |
| Haochen Shi, Shaobo Li, Guoqing Chao, Xiaoliang Shi, Wentao Chen and Zhenzhou Ji                                                                                                                 | 29032 |
| <i>Multilingual vs Crosslingual Retrieval of Fact-Checked Claims: A Tale of Two Approaches</i>                                                                                                   |       |
| Alan Ramponi, Marco Rovera, Robert Moro and Sara Tonelli                                                                                                                                         | 29046 |
| <i>How Much Do LLMs Hallucinate across Languages? On Realistic Multilingual Estimation of LLM Hallucination</i>                                                                                  |       |
| Saad Obaid Ul Islam, Anne Lauscher and Goran Glavaš                                                                                                                                              | 29066 |
| <i>LiTransProQA: An LLM-based Literary Translation Evaluation Metric with Professional Question Answering</i>                                                                                    |       |
| Ran Zhang, Wei Zhao, Lieve Macken and Steffen Eger                                                                                                                                               | 29088 |
| <i>Improving Handshape Representations for Sign Language Processing: A Graph Neural Network Approach</i>                                                                                         |       |
| Alessa Carbo and Eric Nalisnick                                                                                                                                                                  | 29111 |
| <i>Instructing Large Language Models for Low-Resource Languages: A Systematic Study for Basque</i>                                                                                               |       |
| Oscar Sainz, Naiara Perez, Julen Etxaniz, Joseba Fernandez de Landa, Itziar Aldabe, Iker García-Ferrero, Aimar Zabala, Ekhi Azurmendi, German Rigau, Eneko Agirre, Mikel Artetxe and Aitor Soroa | 29125 |
| <i>SOCIAL SCAFFOLDS: A Generalization Framework for Social Understanding Tasks</i>                                                                                                               |       |
| Ritam Dutt, Carolyn Rose and Maarten Sap                                                                                                                                                         | 29150 |
| <i>Beyond A Single AI Cluster: A Survey of Decentralized LLM Training</i>                                                                                                                        |       |
| Haotian Dong, Jingyan Jiang, Rongwei Lu, Jiajun Luo, Jiajun Song, Bowen Li, Ying Shen and Zhi Wang                                                                                               | 29187 |
| <i>Can LLM Agents Maintain a Persona in Discourse?</i>                                                                                                                                           |       |
| Pranav Bhandari, Nicolas Fay, Michael J Wise, Amitava Datta, Stephanie Meek, Usman Naseem and Mehwish Nasim                                                                                      | 29202 |
| <i>Iterative Multilingual Spectral Attribute Erasure</i>                                                                                                                                         |       |
| Shun Shao, Yftah Ziser, Zheng Zhao, Yifu Qiu, Shay B Cohen and Anna Korhonen                                                                                                                     | 29219 |
| <i>TinySQL: A Progressive Text-to-SQL Dataset for Mechanistic Interpretability Research</i>                                                                                                      |       |
| Abir Harrasse, Philip Quirke, Clement Neo, Dhruv Nathawani, Luke Marks and Amir Abdullah                                                                                                         | 29245 |

|                                                                                                                                                                                                                                                         |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>SCRIBE: Structured Chain Reasoning for Interactive Behaviour Explanations using Tool Calling</i><br>Fares Fawzi, Vinitra Swamy, Dominik Glandorf, Tanya Nazaretsky and Tanja Käser . . . . .                                                         | 29274 |
| <i>Logit Space Constrained Fine-Tuning for Mitigating Hallucinations in LLM-Based Recommender Systems</i><br>Jianfeng Deng, Qingfeng Chen, Debo Cheng, Jiuyong Li and Lin Liu . . . . .                                                                 | 29300 |
| <i>PACHAT: Persona-Aware Speech Assistant for Multi-party Dialogue</i><br>Dongjie Fu, Xize Cheng, Linjun Li, Xiaoda Yang, Lujia Yang and Tao Jin . . . . .                                                                                              | 29314 |
| <i>Reasoning-to-Defend: Safety-Aware Reasoning Can Defend Large Language Models from Jailbreaking</i><br>Junda Zhu, Lingyong Yan, Shuaiqiang Wang, Dawei Yin and Lei Sha . . . . .                                                                      | 29332 |
| <i>Graph-Guided Textual Explanation Generation Framework</i><br>Shuzhou Yuan, Jingyi Sun, Ran Zhang, Michael Färber, Steffen Eger, Pepa Atanasova and Isabelle Augenstein . . . . .                                                                     | 29351 |
| <i>The Validation Gap: A Mechanistic Analysis of How Language Models Compute Arithmetic but Fail to Validate It</i><br>Leonardo Bertolazzi, Philipp Mondorf, Barbara Plank and Raffaella Bernardi . . . . .                                             | 29376 |
| <i>A Causal Lens for Evaluating Faithfulness Metrics</i><br>Kerem Zaman and Shashank Srivastava . . . . .                                                                                                                                               | 29414 |
| <i>Sequential-NIAH: A Needle-In-A-Haystack Benchmark for Extracting Sequential Needles from Long Contexts</i><br>Yifei Yu, Qian-Wen Zhang, Lingfeng Qiao, di Yin, Fang Li, Jie Wang, Chen Zeng Xi, Suncong Zheng, Xiaolong Liang and Xing Sun . . . . . | 29439 |
| <i>FISTAPruner: Layer-wise Post-training Pruning for Large Language Models</i><br>Pengxiang Zhao, Hanyu Hu, Ping Li, Yi Zheng, Zhefeng Wang and Xiaoming Yuan . . . . .                                                                                 | 29458 |
| <i>Do LLMs Encode Frame Semantics? Evidence from Frame Identification</i><br>Jayanth Krishna Chundru, Rudrashis Poddar, Jie Cao and Tianyu Jiang . . . . .                                                                                              | 29477 |
| <i>StepER: Step-wise Knowledge Distillation for Enhancing Reasoning Ability in Multi-Step Retrieval-Augmented Language Models</i><br>Kyumin Lee, Minjin Jeon, Sanghwan Jang and Hwanjo Yu . . . . .                                                     | 29490 |
| <i>How Does DPO Reduce Toxicity? A Mechanistic Neuron-Level Analysis</i><br>Yushi Yang, Filip Sondej, Harry Mayne, Andrew Lee and Adam Mahdi . . . . .                                                                                                  | 29513 |
| <i>It's All About In-Context Learning! Teaching Extremely Low-Resource Languages to LLMs</i><br>Yue Li, Zhixue Zhao and Carolina Scarton . . . . .                                                                                                      | 29533 |
| <i>Where to show Demos in Your Prompt: A Positional Bias of In-Context Learning</i><br>Kwesi Adu Cobbina and Tianyi Zhou . . . . .                                                                                                                      | 29549 |
| <i>Multilingual Pretraining for Pixel Language Models</i><br>Ilker Kesen, Jonas F. Lotz, Ingo Ziegler, Phillip Rust and Desmond Elliott . . . . .                                                                                                       | 29583 |
| <i>MetaFaith: Faithful Natural Language Uncertainty Expression in LLMs</i><br>Gabrielle Kaili-May Liu, Gal Yona, Avi Caciularu, Idan Szpektor, Tim G. J. Rudner and Arman Cohan . . . . .                                                               | 29601 |
| <i>Machine-generated text detection prevents language model collapse</i><br>George Drayson, Emine Yilmaz and Vasileios Lamos . . . . .                                                                                                                  | 29646 |

|                                                                                                                                                                                                                                                                                           |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Data-Efficient Hate Speech Detection via Cross-Lingual Nearest Neighbor Retrieval with Limited Labeled Data</i>                                                                                                                                                                        |       |
| Faeze Ghorbanpour, Daryna Dementieva and Alexander Fraser                                                                                                                                                                                                                                 | 29663 |
| <i>V-VAE: A Variational Auto Encoding Framework Towards Fine-Grained Control over Human-Like Chat</i>                                                                                                                                                                                     |       |
| Qi Lin, Weikai Xu, Lisi Chen and Bin Dai                                                                                                                                                                                                                                                  | 29682 |
| <i>Mixture of Languages: Improved Multilingual Encoders Through Language Grouping</i>                                                                                                                                                                                                     |       |
| João Maria Janeiro, Belen Alastruey, Francisco Massa, Maha Elbayad, Benjamin Piwowarski, Patrick Gallinari and Loic Barrault                                                                                                                                                              | 29696 |
| <i>Too Helpful, Too Harmless, Too Honest or Just Right?</i>                                                                                                                                                                                                                               |       |
| Gautam Siddharth Kashyap, Mark Dras and Usman Naseem                                                                                                                                                                                                                                      | 29712 |
| <i>Cardiverse: Harnessing LLMs for Novel Card Game Prototyping</i>                                                                                                                                                                                                                        |       |
| Danrui Li, Sen Zhang, Samuel S. Sohn, Kaidong Hu, Muhammad Usman and Mubbasir Kapadia                                                                                                                                                                                                     | 29724 |
| <i>Assessing effective de-escalation of crisis conversations using transformer-based models and trend statistics</i>                                                                                                                                                                      |       |
| Ignacio J. Tripodi, Greg Buda, Margaret Meagher and Elizabeth A. Olson                                                                                                                                                                                                                    | 29752 |
| <i>Measuring and Mitigating Media Outlet Name Bias in Large Language Models</i>                                                                                                                                                                                                           |       |
| Seong-Jin Park and Kang-Min Kim                                                                                                                                                                                                                                                           | 29767 |
| <i>The Good, the Bad, and the Debatable: A Survey on the Impacts of Data for In-Context Learning</i>                                                                                                                                                                                      |       |
| Stephanie Schoch and Yangfeng Ji                                                                                                                                                                                                                                                          | 29787 |
| <i>Where Confabulation Lives: Latent Feature Discovery in LLMs</i>                                                                                                                                                                                                                        |       |
| Thibaud Ardoin, Yi Cai and Gerhard Wunder                                                                                                                                                                                                                                                 | 29802 |
| <i>Analysing Chain of Thought Dynamics: Active Guidance or Unfaithful Post-hoc Rationalisation?</i>                                                                                                                                                                                       |       |
| Samuel Lewis-Lim, Xingwei Tan, Zhixue Zhao and Nikolaos Aletras                                                                                                                                                                                                                           | 29827 |
| <i>Playpen: An Environment for Exploring Learning From Dialogue Game Feedback</i>                                                                                                                                                                                                         |       |
| Nicola Horst, Davide Mazzaccara, Antonia Schmidt, Michael Sullivan, Filippo Momentè, Luca Franceschetti, Philipp Sadler, Sherzod Hakimov, Alberto Testoni, Raffaella Bernardi, Raquel Fernández, Alexander Koller, Oliver Lemon, David Schlangen, Mario Giulianelli and Alessandro Suglia | 29843 |
| <i>GenLink: Generation-Driven Schema-Linking via Multi-Model Learning for Text-to-SQL</i>                                                                                                                                                                                                 |       |
| Zhifeng Hao, Junqi Huang, Shaobin Shi, Ruichu Cai and Boyan Xu                                                                                                                                                                                                                            | 29881 |
| <i>TSVer: A Benchmark for Fact Verification Against Time-Series Evidence</i>                                                                                                                                                                                                              |       |
| Marek Strong and Andreas Vlachos                                                                                                                                                                                                                                                          | 29895 |
| <i>Cross-MoE: An Efficient Temporal Prediction Framework Integrating Textual Modality</i>                                                                                                                                                                                                 |       |
| Ruizheng Huang, Zhicheng Zhang and Yong Wang                                                                                                                                                                                                                                              | 29916 |
| <i>Sparse Autoencoder Features for Classifications and Transferability</i>                                                                                                                                                                                                                |       |
| Jack Gallifant, Shan Chen, Kuleen Sasse, Hugo Aerts, Thomas Hartvigsen and Danielle Bitterman                                                                                                                                                                                             | 29928 |
| <i>KGE Calibrator: An Efficient Probability Calibration Method of Knowledge Graph Embedding Models for Trustworthy Link Prediction</i>                                                                                                                                                    |       |
| Yang Yang, Mohan Timilsina and Edward Curry                                                                                                                                                                                                                                               | 29953 |
| <i>LCES: Zero-shot Automated Essay Scoring via Pairwise Comparisons Using Large Language Models</i>                                                                                                                                                                                       |       |
| Takumi Shibata and Yuichi Miyamura                                                                                                                                                                                                                                                        | 29977 |

|                                                                                                                                                                               |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>The Arabic Generality Score: Another Dimension of Modeling Arabic Dialectness</i>                                                                                          |       |
| Sanad Sha'ban and Nizar Habash .....                                                                                                                                          | 29991 |
| <i>Lemmatization as a Classification Task: Results from Arabic across Multiple Genres</i>                                                                                     |       |
| Mostafa Saeed and Nizar Habash .....                                                                                                                                          | 30003 |
| <i>A Comprehensive Framework to Operationalize Social Stereotypes for Responsible AI Evaluations</i>                                                                          |       |
| Aida Mostafazadeh Davani, Sunipa Dev, Héctor Pérez-Urbina and Vinodkumar Prabhakaran                                                                                          |       |
| 30019                                                                                                                                                                         |       |
| <i>Correct-Detect: Balancing Performance and Ambiguity Through the Lens of Coreference Resolution in LLMs</i>                                                                 |       |
| Amber Shore, Russell Scheinberg, Ameeta Agrawal and So Young Lee .....                                                                                                        | 30033 |
| <i>GRAID: Synthetic Data Generation with Geometric Constraints and Multi-Agentic Reflection for Harmful Content Detection</i>                                                 |       |
| Melissa Kazemi Rad, Alberto Purpura, Himanshu Kumar, Emily Chen and Mohammad Shahed Sorower .....                                                                             | 30048 |
| <i>LaMDAgent: An Autonomous Framework for Post-Training Pipeline Optimization via LLM Agents</i>                                                                              |       |
| Taro Yano, Yoichi Ishibashi and Masafumi Oyamada .....                                                                                                                        | 30067 |
| <i>Finetuning LLMs for Human Behavior Prediction in Social Science Experiments</i>                                                                                            |       |
| Akaash Kolluri, Shengguang Wu, Joon Sung Park and Michael S. Bernstein .....                                                                                                  | 30085 |
| <i>How Private are Language Models in Abstractive Summarization?</i>                                                                                                          |       |
| Anthony Hughes, Nikolaos Aletras and Ning Ma .....                                                                                                                            | 30101 |
| <i>Expectation Preference Optimization: Reliable Preference Estimation for Improving the Reasoning Capability of Large Language Models</i>                                    |       |
| Zelin Li and Dawei Song .....                                                                                                                                                 | 30120 |
| <i>Split-Merge: Scalable and Memory-Efficient Merging of Expert LLMs</i>                                                                                                      |       |
| Sruthi Gorantla, Aditya Rawal, Devamanyu Hazarika, Kaixiang Lin, Mingyi Hong and Mahdi Namazifar .....                                                                        | 30136 |
| <i>Model Consistency as a Cheap yet Predictive Proxy for LLM Elo Scores</i>                                                                                                   |       |
| Ashwin Ramaswamy, Nestor Demeure and Eral Rrapaj .....                                                                                                                        | 30156 |
| <i>Plutus: Benchmarking Large Language Models in Low-Resource Greek Finance</i>                                                                                               |       |
| Xueqing Peng, Triantafillos Papadopoulos, Efstathia Soufleri, Polydoros Giannouris, Ruoyu Xiang, Yan Wang, Lingfei Qian, Jimin Huang, Qianqian Xie and Sophia Ananiadou ..... | 30165 |
| <i>TaxoAlign: Scholarly Taxonomy Generation Using Language Models</i>                                                                                                         |       |
| Avishek Lahiri, Yufang Hou and Debarshi Kumar Sanyal .....                                                                                                                    | 30192 |
| <i>DiNaM: Disinformation Narrative Mining with Large Language Models</i>                                                                                                      |       |
| Witold Sosnowski, Arkadiusz Modzelewski, Kinga Skorupska and Adam Wierzbicki .....                                                                                            | 30213 |
| <i>VeriLocc: End-to-End Cross-Architecture Register Allocation via LLM</i>                                                                                                    |       |
| Lesheng Jin, Zhenyuan Ruan, Haohui Mai and Jingbo Shang .....                                                                                                                 | 30241 |
| <i>MemeIntel: Explainable Detection of Propagandistic and Hateful Memes</i>                                                                                                   |       |
| Mohamed Bayan Kmainasi, Abul Hasnat, Md Arid Hasan, Ali Ezzat Shahroor and Firoj Alam                                                                                         |       |
| 30252                                                                                                                                                                         |       |

|                                                                                                                                                                                                                                                           |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>FLUID QA: A Multilingual Benchmark for Figurative Language Usage in Dialogue across English, Chinese, and Korean</i>                                                                                                                                   |       |
| Seoyoon Park, Hyeji Choi, Minseon Kim, Subin An, Xiaonan Wang, Gyuri Choi and Hansaem Kim .....                                                                                                                                                           | 30269 |
| <i>Structured Moral Reasoning in Language Models: A Value-Grounded Evaluation Framework</i>                                                                                                                                                               |       |
| Mohna Chakraborty, Lu Wang and David Jurgens .....                                                                                                                                                                                                        | 30284 |
| <i>VerIF: Verification Engineering for Reinforcement Learning in Instruction Following</i>                                                                                                                                                                |       |
| Hao Peng, Yunjia Qi, Xiaozhi Wang, Bin Xu, Lei Hou and Juanzi Li .....                                                                                                                                                                                    | 30313 |
| <i>UNCLE: Benchmarking Uncertainty Expressions in Long-Form Generation</i>                                                                                                                                                                                |       |
| Ruihan Yang, Caiqi Zhang, Zhisong Zhang, Xinting Huang, Dong Yu, Nigel Collier and Deqing Yang .....                                                                                                                                                      | 30329 |
| <i>Enhancing Study-Level Inference from Clinical Trial Papers via Reinforcement Learning-Based Numeric Reasoning</i>                                                                                                                                      |       |
| Massimiliano Pronesti, Michela Lorandi, Paul Flanagan, Oisín Redmond, Anya Belz and Yufang Hou .....                                                                                                                                                      | 30346 |
| <i>Context-aware Biases for Length Extrapolation</i>                                                                                                                                                                                                      |       |
| Ali Veisi, Hamidreza Amirzadeh and Amir M. Mansourian .....                                                                                                                                                                                               | 30363 |
| <i>AutoSDT: Scaling Data-Driven Discovery Tasks Toward Open Co-Scientists</i>                                                                                                                                                                             |       |
| Yifei Li, Hanane Nour Moussa, Ziru Chen, Shijie Chen, Botao Yu, Mingyi Xue, Benjamin Burns, Tzu-Yao Chiu, Vishal Dey, Zitong Lu, Chen Wei, Qianheng Zhang, Tianyu Zhang, Song Gao, Xuhui Huang, Xia Ning, Nesreen K. Ahmed, Ali Payani and Huan Sun ..... | 30385 |
| <i>Finding your MUSE: Mining Unexpected Solutions Engine</i>                                                                                                                                                                                              |       |
| Nir Sweed, Hanit Hakim, Ben Wolfson, Hila Lifshitz and Dafna Shahaf .....                                                                                                                                                                                 | 30408 |
| <i>Quantized but Deceptive? A Multi-Dimensional Truthfulness Evaluation of Quantized LLMs</i>                                                                                                                                                             |       |
| Yao Fu, Xianxuan Long, Runchao Li, Haotian Yu, Mu Sheng, Xiaotian Han, Yu Yin and Pan Li                                                                                                                                                                  | 30424 |
| <i>Leveraging Knowledge Graph-Enhanced LLMs for Context-Aware Medical Consultation</i>                                                                                                                                                                    |       |
| Su-Hyeong Park, Ho-Beom Kim, Seong-Jin Park, Dinara Aliyeva and Kang-Min Kim ....                                                                                                                                                                         | 30448 |
| <i>Reflective Agreement: Combining Self-Mixture of Agents with a Sequence Tagger for Robust Event Extraction</i>                                                                                                                                          |       |
| Fatemeh Haji, Mazal Bethany, Cho-Yu Jason Chiang, Anthony Rios and Peyman Najafirad                                                                                                                                                                       | 30465 |
| <i>Simple Yet Effective: An Information-Theoretic Approach to Multi-LLM Uncertainty Quantification</i>                                                                                                                                                    |       |
| Maya Kruse, Majid Afshar, Saksham Khatwani, Anoop Mayampurath, Guanhua Chen and Yanjun Gao .....                                                                                                                                                          | 30482 |
| <i>Exploring morphology-aware tokenization: A case study on Spanish language modeling</i>                                                                                                                                                                 |       |
| Alba Táboas García, Piotr Przybyła and Leo Wanner .....                                                                                                                                                                                                   | 30494 |
| <i>Studying Rhetorically Ambiguous Questions</i>                                                                                                                                                                                                          |       |
| Oghenevovwe Ikumariogbe, Eduardo Blanco and Ellen Riloff .....                                                                                                                                                                                            | 30508 |
| <i>Estimating LLM Consistency: A User Baseline vs Surrogate Metrics</i>                                                                                                                                                                                   |       |
| Xiaoyuan Wu, Weiran Lin, Omer Akgul and Lujo Bauer .....                                                                                                                                                                                                  | 30519 |
| <i>Are Vision-Language Models Safe in the Wild? A Meme-Based Benchmark Study</i>                                                                                                                                                                          |       |
| DongGeon Lee, Joonwon Jang, Jihae Jeong and Hwanjo Yu .....                                                                                                                                                                                               | 30534 |

|                                                                                                                                                 |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Improving Rule-based Reasoning in LLMs using Neurosymbolic Representations</i>                                                               |       |
| Varun Dhanraj and Chris Eliasmith .....                                                                                                         | 30578 |
| <i>Can LLMs Extract Frame-Semantic Arguments?</i>                                                                                               |       |
| Jacob Devasier, Rishabh Mediratta and Chengkai Li .....                                                                                         | 30598 |
| <i>Accelerated Test-Time Scaling with Model-Free Speculative Sampling</i>                                                                       |       |
| Woomin Song, Saket Dingliwal, Sai Muralidhar Jayanthi, Bhavana Ganesh, Jinwoo Shin, Aram Galstyan and Sravan Babu Bodapati .....                | 30612 |
| <i>Enhancing RLHF with Human Gaze Modeling</i>                                                                                                  |       |
| Karim Galliamov, Ivan Titov and Ilya Pershin .....                                                                                              | 30626 |
| <i>Mapping semantic networks to Dutch word embeddings as a diagnostic tool for cognitive decline</i>                                            |       |
| Maithe van Noort, Michal Korenar and Jelke Bloem .....                                                                                          | 30633 |
| <i>CausalVLBench: Benchmarking Visual Causal Reasoning in Large Vision-Language Models</i>                                                      |       |
| Aneesh Komanduri, Karuna Bhaila and Xintao Wu .....                                                                                             | 30649 |
| <i>Implicit Behavioral Alignment of Language Agents in High-Stakes Crowd Simulations</i>                                                        |       |
| Yunzhe Wang, Gale Lucas, Burcin Becerik-Gerber and Volkan Ustun .....                                                                           | 30670 |
| <i>Are Language Models Consequentialist or Deontological Moral Reasoners?</i>                                                                   |       |
| Keenan Samway, Max Kleiman-Weiner, David Guzman Piedrahita, Rada Mihalcea, Bernhard Schölkopf and Zhijing Jin .....                             | 30688 |
| <i>PatentScore: Multi-dimensional Evaluation of LLM-Generated Patent Claims</i>                                                                 |       |
| Yongmin Yoo, Qionghai Xu and Longbing Cao .....                                                                                                 | 30716 |
| <i>All for One: LLMs Solve Mental Math at the Last Token With Information Transferred From Other Tokens</i>                                     |       |
| Siddarth Mamidanna, Daking Rai, Ziyu Yao and Yilun Zhou .....                                                                                   | 30736 |
| <i>A Position Paper on the Automatic Generation of Machine Learning Leaderboards</i>                                                            |       |
| Roelien C. Timmer, Yufang Hou and Stephen Wan .....                                                                                             | 30750 |
| <i>SimMark: A Robust Sentence-Level Similarity-Based Watermarking Algorithm for Large Language Models</i>                                       |       |
| Amirhossein Dabiriaghdam and Lele Wang .....                                                                                                    | 30774 |
| <i>SERVAL: Surprisingly Effective Zero-Shot Visual Document Retrieval Powered by Large Vision and Language Models</i>                           |       |
| Thong Nguyen, Yibin Lei, Jia-Huei Ju and Andrew Yates .....                                                                                     | 30796 |
| <i>Meta-Semantics Augmented Few-Shot Relational Learning</i>                                                                                    |       |
| Han Wu and Jie Yin .....                                                                                                                        | 30812 |
| <i>ProLongVid: A Simple but Strong Baseline for Long-context Video Instruction Tuning</i>                                                       |       |
| Rui Wang, Bohao Li, Xiyang Dai, Jianwei Yang, Yi-Ling Chen, Zhen Xing, Yifan Yang, Dongdong Chen, Xipeng Qiu, Zuxuan Wu and Yu-Gang Jiang ..... | 30825 |
| <i>ModelCitizens: Representing Community Voices in Online Safety</i>                                                                            |       |
| Ashima Suvana, Christina A Chance, Karolina Naranjo, Hamid Palangi, Sophie Hao, Thomas Hartvigsen and Saadia Gabriel .....                      | 30839 |
| <i>UnifiedVisual: A Framework for Constructing Unified Vision-Language Datasets</i>                                                             |       |
| Pengyu Wang, Shaojun Zhou, Chenkun Tan, Xinghao Wang, Wei Huang, Zhen Ye, Zhaowei Li, Botian Jiang, Dong Zhang and Xipeng Qiu .....             | 30856 |



|                                                                                                                                   |       |
|-----------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>The Pursuit of Empathy: Evaluating Small Language Models for PTSD Dialogue Support</i>                                         |       |
| Suhas BN, Yash Mahajan, Dominik O. Mattioli, Andrew M. Sherrill, Rosa I. Arriaga, Christopher Wiese and Saeed Abdullah . . . . .  | 30889 |
| <i>Is Cognition Consistent with Perception? Assessing and Mitigating Multimodal Knowledge Conflicts in Document Understanding</i> |       |
| Zirui Shao, Feiyu Gao, Zhaoqing Zhu, Chuwei Luo, Hangdi Xing, Zhi Yu, Qi Zheng, Ming Yan and Jiajun Bu . . . . .                  | 30912 |
| <i>AutoCT: Automating Interpretable Clinical Trial Prediction with LLM Agents</i>                                                 |       |
| Fengze Liu, Haoyu Wang, Joonhyuk Cho, Dan Roth and Andrew Lo . . . . .                                                            | 30934 |
| <i>MMDocIR: Benchmarking Multimodal Retrieval for Long Documents</i>                                                              |       |
| Kuicai Dong, Yujing Chang, Derrick Goh Xin Deik, Dexun Li, Ruiming Tang and Yong Liu                                              | 30960 |
| <i>Program of Thoughts for Financial Reasoning: Leveraging Dynamic In-Context Examples and Generative Retrieval</i>               |       |
| Subhendu Khatuya, Shashwat Naidu, Pawan Goyal and Niloy Ganguly . . . . .                                                         | 30995 |
| <i>Waste-Bench: A Comprehensive Benchmark for Evaluating VLLMs in Cluttered Environments</i>                                      |       |
| Muhammad Ali and Salman Khan . . . . .                                                                                            | 31008 |
| <i>Demystifying Domain-adaptive Post-training for Financial LLMs</i>                                                              |       |
| Zixuan Ke, Yifei Ming, Xuan-Phi Nguyen, Caiming Xiong and Shafiq Joty . . . . .                                                   | 31022 |
| <i>HICode: Hierarchical Inductive Coding with LLMs</i>                                                                            |       |
| Mian Zhong, Pristina Wang and Anjalie Field . . . . .                                                                             | 31049 |
| <i>Cacheback: Speculative Decoding With Nothing But Cache</i>                                                                     |       |
| Zhiyao Ma, In Gim and Lin Zhong . . . . .                                                                                         | 31068 |
| <i>MA-DPR: Manifold-aware Distance Metrics for Dense Passage Retrieval</i>                                                        |       |
| Yifan Liu, Qianfeng Wen, Mark Zhao, Jiazhou Liang and Scott Sanner . . . . .                                                      | 31074 |
| <i>LLM-Guided Co-Training for Text Classification</i>                                                                             |       |
| Md Mezbaur Rahman and Cornelia Caragea . . . . .                                                                                  | 31093 |
| <i>LeanK: Learnable K Cache Channel Pruning for Efficient Decoding</i>                                                            |       |
| Yike Zhang, Zhiyuan He, Huiqiang Jiang, Chengruidong Zhang, Yuqing Yang, Jianyong Wang and Lili Qiu . . . . .                     | 31111 |
| <i>DELOC: Document Element Localizer</i>                                                                                          |       |
| Hammad Ayyubi, Puneet Mathur, Mehrab Tanjim and Vlad I Morariu . . . . .                                                          | 31127 |
| <i>NL2Lean: Translating Natural Language into Lean 4 through Multi-Aspect Reinforcement Learning</i>                              |       |
| Yue Fang, Shaohan Huang, Xin Yu, Haizhen Huang, Zihan Zhang, Weiwei Deng, Furu Wei, Feng Sun, Qi Zhang and Zhi Jin . . . . .      | 31137 |
| <i>A Multilingual, Culture-First Approach to Addressing Misgendering in LLM Applications</i>                                      |       |
| Sunayana Sitaram, Adrian de Wynter, Isobel McCrum, Qilong Gu and Si-Qing Chen . . . . .                                           | 31148 |
| <i>X-CoT: Explainable Text-to-Video Retrieval via LLM-based Chain-of-Thought Reasoning</i>                                        |       |
| Prasanna Reddy Pulakurthi, Jiamian Wang, Majid Rabbani, Sohail Dianat, Raghuveer Rao and Zhiqiang Tao . . . . .                   | 31173 |

|                                                                                                                                                                                                            |       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Token-level Proximal Policy Optimization for Query Generation</i>                                                                                                                                       |       |
| Yichen Ouyang, Lu Wang, Fangkai Yang, Pu Zhao, Chenghua Huang, Jianfeng Liu, Bochen Pang, Yaming Yang, Yuefeng Zhan, Hao Sun, Qingwei Lin, Saravan Rajmohan, Weiwei Deng, Dongmei Zhang and Feng Sun ..... | 31185 |
| <i>Prior Prompt Engineering for Reinforcement Fine-Tuning</i>                                                                                                                                              |       |
| Pittawat Taveekitworachai, Potsawee Manakul, Sarana Nutanong and Kunat Pipatanakul ..                                                                                                                      | 31200 |
| <i>Beyond WER: Probing Whisper’s Sub-token Decoder Across Diverse Language Resource Levels</i>                                                                                                             |       |
| Siyu Liang, Nicolas Ballier, Gina-Anne Levow and Richard Wright .....                                                                                                                                      | 31226 |
| <i>ThinkTuning: Instilling Cognitive Reflections without Distillation</i>                                                                                                                                  |       |
| Aswin Rrv, Jacob Dineen, Divij Handa, Md Nayem Uddin, Mihir Parmar, Chitta Baral and Ben Zhou .....                                                                                                        | 31237 |
| <i>Droid: A Resource Suite for AI-Generated Code Detection</i>                                                                                                                                             |       |
| Daniil Orel, Indraneil Paul, Iryna Gurevych and Preslav Nakov .....                                                                                                                                        | 31252 |
| <i>LoRACoE: Improving Large Language Model via Composition-based LoRA Expert</i>                                                                                                                           |       |
| Guanyu Li, Zhiheng Xi, Zhihao Zhang, Boyang Hong, Tao Gui, Qi Zhang and Xuanjing Huang                                                                                                                     | 31279 |
| <i>Same Question, Different Words: A Latent Adversarial Framework for Prompt Robustness</i>                                                                                                                |       |
| Tingchen Fu and Fazl Barez .....                                                                                                                                                                           | 31294 |
| <i>Pluralistic Alignment for Healthcare: A Role-Driven Framework</i>                                                                                                                                       |       |
| Jiayou Zhong, Anudeex Shetty, Chao Jia, Xuanrui Lin and Usman Naseem .....                                                                                                                                 | 31309 |
| <i>Flexible-length Text Infilling for Discrete Diffusion Models</i>                                                                                                                                        |       |
| Andrew Zhang, Anushka Sivakumar, Chia-Wei Tang and Chris Thomas .....                                                                                                                                      | 31333 |
| <i>Beyond the Leaderboard: Understanding Performance Disparities in Large Language Models via Model Diffing</i>                                                                                            |       |
| Sabri Boughorbel, Fahim Dalvi, Nadir Durrani and Majd Hawasly .....                                                                                                                                        | 31349 |
| <i>Explicit Learning and the LLM in Machine Translation</i>                                                                                                                                                |       |
| Malik Marmonier, Rachel Bawden and Benoît Sagot .....                                                                                                                                                      | 31361 |
| <i>Towards Language-Agnostic STIPA: Universal Phonetic Transcription to Support Language Documentation at Scale</i>                                                                                        |       |
| Jacob Lee Suchardt, Hana El-Shazli and Pierluigi Cassotti .....                                                                                                                                            | 31412 |
| <i>Beyond Pairwise: Global Zero-shot Temporal Graph Generation</i>                                                                                                                                         |       |
| Alon Eirew, Kfir Bar and Ido Dagan .....                                                                                                                                                                   | 31429 |
| <i>“Feels Feminine to Me”: Understanding Perceived Gendered Style through Human Annotations</i>                                                                                                            |       |
| Hongyu Chen, Neele Falk, Michael Roth and Agnieszka Falenska .....                                                                                                                                         | 31448 |
| <i>RALS: Resources and Baselines for Romanian Automatic Lexical Simplification</i>                                                                                                                         |       |
| Fabian Anghel, Cristea Petru-Theodor, Claudiu Creanga and Sergiu Nisioi .....                                                                                                                              | 31470 |
| <i>How Do Social Bots Participate in Misinformation Spread? A Comprehensive Dataset and Analysis</i>                                                                                                       |       |
| Herun Wan, Minnan Luo, Zihan Ma, Guang Dai and Xiang Zhao .....                                                                                                                                            | 31482 |
| <i>Are Stereotypes Leading LLMs’ Zero-Shot Stance Detection ?</i>                                                                                                                                          |       |
| Anthony Dubreuil, Antoine Gourru, Christine Largeron and Amine Trabelsi .....                                                                                                                              | 31506 |



|                                                                                                                                     |       |
|-------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Multi-Modal Framing Analysis of News</i>                                                                                         |       |
| Arnav Arora, Srishti Yadav, Maria Antoniak, Serge Belongie and Isabelle Augenstein . . .                                            | 31520 |
| <i>TempParaphraser: Heating UpText to Evade AI-Text Detection through Paraphrasing</i>                                              |       |
| Junjie Huang, Ruiquan Zhang, Jinsong Su and Yidong Chen . . . . .                                                                   | 31543 |
| <i>ComicScene154: A Scene Dataset for Comic Analysis</i>                                                                            |       |
| Sandro Paval, Pascal Meißner and Ivan P. Yamshchikov . . . . .                                                                      | 31563 |
| <i>MedLinkDE – MedDRA Entity Linking for German with Guided Chain of Thought Reasoning</i>                                          |       |
| Roman Christof, Farnaz Zeidi, Manuela Messelhäuser, Dirk Mentzer, Renate Koenig, Liam Childs and Alexander Mehler . . . . .         | 31570 |
| <i>HookMoE: A learnable performance compensation strategy of Mixture-of-Experts for LLM inference acceleration</i>                  |       |
| Cheng Longkai, Along He, Mulin Li, Xie Xueshuo and Tao Li . . . . .                                                                 | 31583 |
| <i>Cross-Document Cross-Lingual NLI via RST-Enhanced Graph Fusion and Interpretability Prediction</i>                               |       |
| Mengying Yuan, WenHao Wang, Zixuan Wang, Yujie Huang, Kangli Wei, Fei Li, Chong Teng and Donghong Ji . . . . .                      | 31596 |
| <i>3R: Enhancing Sentence Representation Learning via Redundant Representation Reduction</i>                                        |       |
| Longxuan Ma, Xiao Wu, Yuxin Huang, Shengxiang Gao and Zhengtao Yu . . . . .                                                         | 31619 |
| <i>When Big Models Train Small Ones: Label-Free Model Parity Alignment for Efficient Visual Question Answering using Small VLMs</i> |       |
| Abhirama Subramanyam Penamakuri, Navlika Singh, Piyush Arora and Anand Mishra . . .                                                 | 31633 |
| <i>ProReason: Multi-Modal Proactive Reasoning with Decoupled Eyesight and Wisdom</i>                                                |       |
| Jingqi Zhou, Sheng Wang, Jingwei Dong, Kai Liu, Lei Li, Jiahui Gao, Jiyue Jiang, Lingpeng Kong and Chuan Wu . . . . .               | 31651 |
| <i>Extractive Fact Decomposition for Interpretable Natural Language Inference in one Forward Pass</i>                               |       |
| Nicholas Popović and Michael Färber . . . . .                                                                                       | 31681 |
| <i>Structure-Conditional Minimum Bayes Risk Decoding</i>                                                                            |       |
| Bryan Eikema, Anna Rutkiewicz and Mario Giulianelli . . . . .                                                                       | 31695 |
| <i>Label Set Optimization via Activation Distribution Kurtosis for Zero-Shot Classification with Generative Models</i>              |       |
| Yue Li, Zhixue Zhao and Carolina Scarton . . . . .                                                                                  | 31713 |
| <i>The Transfer Neurons Hypothesis: An Underlying Mechanism for Language Latent Space Transitions in Multilingual LLMs</i>          |       |
| Hinata Tezuka and Naoya Inoue . . . . .                                                                                             | 31731 |
| <i>VEHME: A Vision-Language Model For Evaluating Handwritten Mathematics Expressions</i>                                            |       |
| Thu Phuong Nguyen, Duc M. Nguyen, Hyotaek Jeon, Hyunwook Lee, Hyunmin Song, Sungahn Ko and Taehwan Kim . . . . .                    | 31782 |
| <i>All Roads Lead to Rome: Graph-Based Confidence Estimation for Large Language Model Reasoning</i>                                 |       |
| Caiqi Zhang, Chang Shu, Ehsan Shareghi and Nigel Collier . . . . .                                                                  | 31803 |
| <i>SEMMA: A Semantic Aware Knowledge Graph Foundation Model</i>                                                                     |       |
| Arvinth Arun, Sumit Kumar, Mojtaba Nayyeri, Bo Xiong, Ponnurangam Kumaraguru, Antonio Vergari and Steffen Staab . . . . .           | 31814 |

|                                                                                                                                                                                                                                        |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Text2Vis: A Challenging and Diverse Benchmark for Generating Multimodal Visualizations from Text</i><br>Mizanur Rahman, Md Tahmid Rahman Laskar, Shafiq Joty and Enamul Hoque .....                                                 | 31838 |
| <i>Predicting Prosodic Boundaries for Children's Texts</i><br>Mansi Dhamne, Sneha Raman and Preeti Rao .....                                                                                                                           | 31864 |
| <i>Enhancing Logical Reasoning in Language Models via Symbolically-Guided Monte Carlo Process Supervision</i><br>Xingwei Tan, Marco Valentino, Mahmud Elahi Akhter, Maria Liakata and Nikolaos Aletras                                 | 31875 |
| <i>Can Large Language Models Outperform Non-Experts in Poetry Evaluation? A Comparative Study Using the Consensual Assessment Technique</i><br>Piotr Sawicki, Marek Grzes, Dan Brown and Fabricio Goes .....                           | 31890 |
| <i>Beyond Human Labels: A Multi-Linguistic Auto-Generated Benchmark for Evaluating Large Language Models on Resume Parsing</i><br>Zijian Ling, Han Zhang, Jiahao Cui, Zhequn Wu, Xu Sun, Guohao Li and Xiangjian He ..                 | 31908 |
| <i>Orthogonal Finetuning Made Scalable</i><br>Zeju Qiu, Weiyang Liu, Adrian Weller and Bernhard Schölkopf .....                                                                                                                        | 31935 |
| <i>AIR: Complex Instruction Generation via Automatic Iterative Refinement</i><br>Wei Liu, Yancheng He, Yu Li, Hui Huang, Chengwei Hu, Jiaheng Liu, Shilong Li, Wenbo Su and Bo Zheng .....                                             | 31953 |
| <i>SQUiD: Synthesizing Relational Databases from Unstructured Text</i><br>Mushtari Sadia, Zhenning Yang, Yunming Xiao, Ang Chen and Amrita Roy Chowdhury ..                                                                            | 31976 |
| <i>RAG+: Enhancing Retrieval-Augmented Generation with Application-Aware Reasoning</i><br>Yu Wang, Shiwan Zhao, Zhihu Wang, Ming Fan, Xicheng Zhang, Yubo Zhang, Zhengfan Wang, Heyuan Huang and Ting Liu .....                        | 32002 |
| <i>Rapid Word Learning Through Meta In-Context Learning</i><br>Wentao Wang, Guangyuan Jiang, Tal Linzen and Brenden Lake .....                                                                                                         | 32027 |
| <i>EuroGEST: Investigating gender stereotypes in multilingual language models</i><br>Jacqueline Rowe, Mateusz Klimaszewski, Liane Guillou, Shannon Vallor and Alexandra Birch                                                          | 32063 |
| <i>How Persuasive Is Your Context?</i><br>Tu Nguyen, Kevin Du, Alexander Miserlis Hoyle and Ryan Cotterell .....                                                                                                                       | 32086 |
| <i>The Medium Is Not the Message: Deconfounding Document Embeddings via Linear Concept Erasure</i><br>Yu Fan, Yang Tian, Shauli Ravfogel, Mrinmaya Sachan, Elliott Ash and Alexander Miserlis Hoyle                                    | 32113 |
| <i>Measuring scalar constructs in social science with LLMs</i><br>Hauke Licht, Rupak Sarkar, Patrick Y. Wu, Pranav Goel, Niklas Stoeck, Elliott Ash and Alexander Miserlis Hoyle .....                                                 | 32133 |
| <i>Text Detoxification: Data Efficiency, Semantic Preservation and Model Generalization</i><br>Jing Yu, Yibo Zhao, Jiapeng Zhu, Wenming Shao, Bo Pang, Zhao Zhang and Xiang Li ...                                                     | 32161 |
| <i>Not What the Doctor Ordered: Surveying LLM-based De-identification and Quantifying Clinical Information Loss</i><br>Kiana Aghakasiri, Noopur Zambare, JoAnn Thai, Carrie Ye, Mayur Mehta, J Ross Mitchell and Mohamed Abdalla ..... | 32176 |

|                                                                                                                                                                                |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Reasoning under Uncertainty: Efficient LLM Inference via Unsupervised Confidence Dilution and Convergent Adaptive Sampling</i>                                              |       |
| Zhenning Shi, Yijia Zhu, Yi Xie, Junhan Shi, Guorui Xie, Haotian Zhang, Yong Jiang, Congcong Miao and Qing Li . . . . .                                                        | 32193 |
| <i>Africa Health Check: Probing Cultural Bias in Medical LLMs</i>                                                                                                              |       |
| Charles Nimo, Shuheng Liu, Irfan Essa and Michael L. Best . . . . .                                                                                                            | 32208 |
| <i>Assumed Identities: Quantifying Gender Bias in Machine Translation of Gender-Ambiguous Occupational Terms</i>                                                               |       |
| Orfeas Menis Mastromichalakis, Giorgos Filandrianos, Maria Symeonaki and Giorgos Stamou . . . . .                                                                              | 32222 |
| <i>REVIVING YOUR MNEME: Predicting The Side Effects of LLM Unlearning and Fine-Tuning via Sparse Model Diffing</i>                                                             |       |
| Aly M. Kassem, Zhuan Shi, Negar Rostamzadeh and Golnoosh Farnadi . . . . .                                                                                                     | 32239 |
| <i>ToM-SSI: Evaluating Theory of Mind in Situated Social Interactions</i>                                                                                                      |       |
| Matteo Bortoletto, Constantin Ruhdorfer and Andreas Bulling . . . . .                                                                                                          | 32253 |
| <i>Recursive Training Loops in LLMs: How training data properties modulate distribution shift in generated data?</i>                                                           |       |
| Grgur Kovač, Jérémy Perez, Rémy Portelas, Peter Ford Dominey and Pierre-Yves Oudeyer . . . . .                                                                                 | 32279 |
| <i>Detecting LLM Hallucination Through Layer-wise Information Deficiency: Analysis of Ambiguous Prompts and Unanswerable Questions</i>                                         |       |
| Hazel Kim, Tom A. Lamb, Adel Bibi, Philip Torr and Yarin Gal . . . . .                                                                                                         | 32299 |
| <i>Extending Automatic Machine Translation Evaluation to Book-Length Documents</i>                                                                                             |       |
| Kuang-Da Wang, Shuoyang Ding, Chao-Han Huck Yang, Ping-Chun Hsieh, Wen-Chih Peng, Vitaly Lavrukhin and Boris Ginsburg . . . . .                                                | 32312 |
| <i>MedFact: A Large-scale Chinese Dataset for Evidence-based Medical Fact-checking of LLM Responses</i>                                                                        |       |
| Tong Chen, Zimu Wang, Yiyi Miao, Haoran Luo, Sun Yuanfei, Wei Wang, Zhengyong Jiang, Procheta Sen and Jionglong Su . . . . .                                                   | 32329 |
| <i>VideoPASTA: 7K Preference Pairs That Matter for Video-LLM Alignment</i>                                                                                                     |       |
| Yogesh Kulkarni and Pooyan Fazli . . . . .                                                                                                                                     | 32343 |
| <i>Do LLMs Adhere to Label Definitions? Examining Their Receptivity to External Label Definitions</i>                                                                          |       |
| Seyedali Mohammadi, Bhaskara Hanuma Vedula, Hemank Lamba, Edward Raff, Ponnurangam Kumaraguru, Francis Ferraro and Manas Gaur . . . . .                                        | 32369 |
| <i>Group-Aware Reinforcement Learning for Output Diversity in Large Language Models</i>                                                                                        |       |
| Oron Anschel, Alon Shoshan, Adam Botach, Shunit Haviv Hakimi, Asaf Gendler, Emanuel Ben Baruch, Nadav Bhonker, Igor Kviatkovsky, Manoj Aggarwal and Gerard Medioni . . . . .   | 32383 |
| <i>Model-Based Ranking of Source Languages for Zero-Shot Cross-Lingual Transfer</i>                                                                                            |       |
| Abteen Ebrahimi, Adam Wiemerslage and Katharina von der Wense . . . . .                                                                                                        | 32405 |
| <i>PruneCD: Contrasting Pruned Self Model to Improve Decoding Factuality</i>                                                                                                   |       |
| Byeongho Yu, Changhun Lee, Jun-gyu Jin and Eunhyeok Park . . . . .                                                                                                             | 32451 |
| <i>Crisp: Cognitive Restructuring of Negative Thoughts through Multi-turn Supportive Dialogues</i>                                                                             |       |
| Jinfeng Zhou, Yuxuan Chen, Jianing Yin, Yongkang Huang, Yihan Shi, Xikun Zhang, Libiao Peng, Rongsheng Zhang, Tangjie Lv, Zhipeng Hu, Hongning Wang and Minlie Huang . . . . . | 32463 |

|                                                                                                                                                    |       |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>AccessEval: Benchmarking Disability Bias in Large Language Models</i>                                                                           |       |
| Srikant Panda, Amit Agarwal and Hitesh Laxmichand Patel .....                                                                                      | 32493 |
| <i>The Impact of Language Mixing on Bilingual LLM Reasoning</i>                                                                                    |       |
| Yihao Li, Jiayi Xin, Miranda Muqing Miao, Qi Long and Lyle Ungar .....                                                                             | 32520 |
| <i>VISaGE: Understanding Visual Generics and Exceptions</i>                                                                                        |       |
| Stella Frank and Emily Allaway .....                                                                                                               | 32538 |
| <i>Stronger Baselines for Retrieval-Augmented Generation with Long-Context Language Models</i>                                                     |       |
| Alex Laitenberger, Christopher D Manning and Nelson F. Liu .....                                                                                   | 32548 |
| <i>Discursive Circuits: How Do Language Models Understand Discourse Relations?</i>                                                                 |       |
| Yisong Miao and Min-Yen Kan .....                                                                                                                  | 32559 |
| <i>Making VLMs More Robot-Friendly: Self-Critical Distillation of Low-Level Procedural Reasoning</i>                                               |       |
| Chan Young Park, Jillian Fisher, Marius Memmel, Dipika Khullar, Seoho Yun, Abhishek Gupta and Yejin Choi .....                                     | 32579 |
| <i>ThinkSLM: Towards Reasoning in Small Language Models</i>                                                                                        |       |
| Gaurav Srivastava, Shuxiang Cao and Xuan Wang .....                                                                                                | 32601 |
| <i>MAGICoRe: Multi-Agent, Iterative, Coarse-to-Fine Refinement for Reasoning</i>                                                                   |       |
| Justin Chen, Archiki Prasad, Swarnadeep Saha, Elias Stengel-Eskin and Mohit Bansal ....                                                            | 32652 |
| <i>Batched Self-Consistency Improves LLM Relevance Assessment and Ranking</i>                                                                      |       |
| Anton Korikov, Pan Du, Scott Sanner and Navid Rekabsaz .....                                                                                       | 32676 |
| <i>SemCSE: Semantic Contrastive Sentence Embeddings Using LLM-Generated Summaries For Scientific Abstracts</i>                                     |       |
| Marc Felix Brinner and Sina Zarri   .....                                                                                                          | 32693 |
| <i>Controlled Generation for Private Synthetic Text</i>                                                                                            |       |
| Zihao Zhao and Anjalie Field .....                                                                                                                 | 32709 |
| <i>Towards AI-Assisted Psychotherapy: Emotion-Guided Generative Interventions</i>                                                                  |       |
| Kilichbek Haydarov, Youssef Mohamed, Emilio Goldenhersch, Paul OCallaghan, Li-jia Li and Mohamed Elhoseiny .....                                   | 32725 |
| <i>From Shortcuts to Balance: Attribution Analysis of Speech-Text Feature Utilization in Distinguishing Original from Machine-Translated Texts</i> |       |
| Yongjian Chen and Antonio Toral .....                                                                                                              | 32745 |
| <i>DEBATE, TRAIN, EVOLVE: Self-Evolution of Language Model Reasoning</i>                                                                           |       |
| Gaurav Srivastava, Zhenyu Bi, Meng Lu and Xuan Wang .....                                                                                          | 32753 |
| <i>From Chat Logs to Collective Insights: Aggregative Question Answering</i>                                                                       |       |
| Wentao Zhang, Woojeong Kim and Yuntian Deng .....                                                                                                  | 32800 |
| <i>A Text-Based Recommender System that Leverages Explicit Affective State Preferences</i>                                                         |       |
| Tonmoy Hasan and Razvan Bunescu .....                                                                                                              | 32840 |
| <i>CARE: Multilingual Human Preference Learning for Cultural Awareness</i>                                                                         |       |
| Geyang Guo, Tarek Naous, Hiromi Wakaki, Yukiko Nishimura, Yuki Mitsufuji, Alan Ritter and Wei Xu .....                                             | 32855 |
| <i>Multilingual Dialogue Generation and Localization with Dialogue Act Scripting</i>                                                               |       |
| Justin Vasselli, Eunike Andriani Kardinata, Yusuke Sakai and Taro Watanabe .....                                                                   | 32885 |

|                                                                                                                                                                                                                                                                                                                                                                 |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>SUE: Sparsity-based Uncertainty Estimation via Sparse Dictionary Learning</i>                                                                                                                                                                                                                                                                                |       |
| Tamás Ficsor and Gábor Berend .....                                                                                                                                                                                                                                                                                                                             | 32901 |
| <i>Planning-Aware Code Infilling via Horizon-Length Prediction</i>                                                                                                                                                                                                                                                                                              |       |
| Yifeng Ding, Hantian Ding, Shiqi Wang, Qing Sun, Varun Kumar and Zijian Wang .....                                                                                                                                                                                                                                                                              | 32919 |
| <i>SinhalaMMLU: A Comprehensive Benchmark for Evaluating Multitask Language Understanding in Sinhala</i>                                                                                                                                                                                                                                                        |       |
| Ashmari Pramodya, Nirasha Nelki, Heshan Shalinda, Chamila Liyanage, Yusuke Sakai, Randil Pushpananda, Ruvan Weerasinghe, Hidetaka Kamigaito and Taro Watanabe .....                                                                                                                                                                                             | 32932 |
| <i>OG-RAG: Ontology-grounded retrieval-augmented generation for large language models</i>                                                                                                                                                                                                                                                                       |       |
| Kartik Sharma, Peeyush Kumar and Yunqing Li .....                                                                                                                                                                                                                                                                                                               | 32951 |
| <i>Convergence and Divergence of Language Models under Different Random Seeds</i>                                                                                                                                                                                                                                                                               |       |
| Finlay Fehlaue, Kyle Mahowald and Tiago Pimentel .....                                                                                                                                                                                                                                                                                                          | 32971 |
| <i>Analyzing and Modeling LLM Response Lengths with Extreme Value Theory: Anchoring Effects and Hybrid Distributions</i>                                                                                                                                                                                                                                        |       |
| Liuxuan Jiao, Chen Gao, Yiqian Yang, Chenliang Zhou, YiXian Huang, Xinlei Chen and Yong Li .....                                                                                                                                                                                                                                                                | 32981 |
| <i>Language Models Identify Ambiguities and Exploit Loopholes</i>                                                                                                                                                                                                                                                                                               |       |
| Jio Choi, Mohit Bansal and Elias Stengel-Eskin .....                                                                                                                                                                                                                                                                                                            | 32992 |
| <i>Benchmarking LLMs for Translating Classical Chinese Poetry: Evaluating Adequacy, Fluency, and Elegance</i>                                                                                                                                                                                                                                                   |       |
| Andong Chen, Lianzhang Lou, Kehai Chen, Xuefeng Bai, Yang Xiang, Muyun Yang, Tiejun Zhao and Min Zhang .....                                                                                                                                                                                                                                                    | 33008 |
| <i>AraEval: An Arabic Multi-Task Evaluation Suite for Large Language Models</i>                                                                                                                                                                                                                                                                                 |       |
| Alhanoof Althnian, Norah A. Alzahrani, Shaykhah Z. Alsubaie, Eman Albilali, Ahmed Abdelali, Nouf M. Alotaibi, M Saiful Bari, Yazeed Alnumay, Abdulhamed Alothaimen, Maryam Saif, Shahad D. Alzaidi, Faisal Abdulrahman Mirza, Yousef Almushayqih, Mohammed Al Saleem, Ghadah Alabduljabbar, Abdulmohsen Al-Thubaity, Areeb Alowisheq and Nora Al-Twairesh ..... | 33026 |
| <i>QUIDS: Query Intent Description for Exploratory Search via Dual Space Modeling</i>                                                                                                                                                                                                                                                                           |       |
| Yumeng Wang, Xiuying Chen and Suzan Verberne .....                                                                                                                                                                                                                                                                                                              | 33051 |
| <i>A Systematic Survey of Automatic Prompt Optimization Techniques</i>                                                                                                                                                                                                                                                                                          |       |
| Kiran Ramnath, Kang Zhou, Sheng Guan, Soumya Smruti Mishra, Xuan Qi, Zhengyuan Shen, Shuai Wang, Sangmin Woo, Sullam Jeoung, Yawei Wang, Haozhu Wang, Han Ding, Yuzhe Lu, Zhichao Xu, Yun Zhou, Balasubramaniam Srinivasan, Qiaojing Yan, Yueyan Chen, Haibo Ding, Panpan Xu and Lin Lee Cheong .....                                                           | 33067 |
| <i>Threading the Needle: Reweaving Chain-of-Thought Reasoning to Explain Human Label Variation</i>                                                                                                                                                                                                                                                              |       |
| Beiduo Chen, Yang Janet Liu, Anna Korhonen and Barbara Plank .....                                                                                                                                                                                                                                                                                              | 33100 |
| <i>MemInsight: Autonomous Memory Augmentation for LLM Agents</i>                                                                                                                                                                                                                                                                                                |       |
| Rana Salama, Jason Cai, Michelle Yuan, Anna Currey, Monica Sunkara, Yi Zhang and Yassine Benajiba .....                                                                                                                                                                                                                                                         | 33125 |
| <i>Breaking the Noise Barrier: LLM-Guided Semantic Filtering and Enhancement for Multi-Modal Entity Alignment</i>                                                                                                                                                                                                                                               |       |
| Chenglong Lu, Chenxiao Li, Jingwei Cheng, Yongquan Ji, Guoqing Chen and Fu Zhang ..                                                                                                                                                                                                                                                                             | 33142 |

|                                                                                                                                                             |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>ImpliRet: Benchmarking the Implicit Fact Retrieval Challenge</i>                                                                                         |       |
| Zeinab Sadat Taghavi, Ali Modarressi, Yunpu Ma and Hinrich Schuetze . . . . .                                                                               | 33157 |
| <i>No Need for Explanations: LLMs can implicitly learn from mistakes in-context</i>                                                                         |       |
| Lisa Alazraki, Maximilian Mozes, Jon Ander Campos, Tan Yi-Chern, Marek Rei and Max Bartolo                                                                  | 33180 |
| <i>MoVa: Towards Generalizable Classification of Human Morals and Values</i>                                                                                |       |
| Ziyu Chen, Junfei Sun, Chenxi Li, Tuan Dung Nguyen, Jing Yao, Xiaoyuan Yi, Xing Xie, Chenhao Tan and Lexing Xie . . . . .                                   | 33205 |
| <i>GUI-Bee: Align GUI Action Grounding to Novel Environments via Autonomous Exploration</i>                                                                 |       |
| Yue Fan, Handong Zhao, Ruiyi Zhang, Yu Shen, Xin Eric Wang and Gang Wu . . . . .                                                                            | 33250 |
| <i>Revealing and Mitigating the Challenge of Detecting Character Knowledge Errors in LLM Role-Playing</i>                                                   |       |
| Wenyuan Zhang, Shuaiyi Nie, Jiawei Sheng, Zefeng Zhang, Xinghua Zhang, Yongquan He and Tingwen Liu . . . . .                                                | 33268 |
| <i>Taking Notes Brings Focus? Towards Multi-Turn Multimodal Dialogue Learning</i>                                                                           |       |
| Jiazheng Liu, Sipeng Zheng, Börje F. Karlsson and Zongqing Lu . . . . .                                                                                     | 33292 |
| <i>Graph-Based Multi-Trait Essay Scoring</i>                                                                                                                |       |
| Shengjie Li and Vincent Ng . . . . .                                                                                                                        | 33314 |
| <i>Benchmarking LLMs on Semantic Overlap Summarization</i>                                                                                                  |       |
| John Salvador, Naman Bansal, Mousumi Akter, Souvika Sarkar, Anupam Das and Santu Karmaker . . . . .                                                         | 33341 |
| <i>N-CORE: N-View Consistency Regularization for Disentangled Representation Learning in Nonverbal Vocalizations</i>                                        |       |
| Siddhant Bikram Shah and Kristina T. Johnson . . . . .                                                                                                      | 33363 |
| <i>Probability Distribution Collapse: A Critical Bottleneck to Compact Unsupervised Neural Grammar Induction</i>                                            |       |
| Jinwook Park and Kangil Kim . . . . .                                                                                                                       | 33381 |
| <i>Spatial Layouts in News Homepages Capture Human Preferences</i>                                                                                          |       |
| Alexander Spangher, Michael Vu, Arda Kaz, Naitian Zhou and Ben Welsh . . . . .                                                                              | 33393 |
| <i>KRETA: A Benchmark for Korean Reading and Reasoning in Text-Rich VQA Attuned to Diverse Visual Contexts</i>                                              |       |
| Taeback Hwang, Minseo Kim, Gisang Lee, Seonuk Kim and Hyunjun Eun . . . . .                                                                                 | 33410 |
| <i>ReflAct: World-Grounded Decision Making in LLM Agents via Goal-State Reflection</i>                                                                      |       |
| Jeonghye Kim, Sojeong Rhee, Minbeom Kim, Dohyung Kim, Sangmook Lee, Youngchul Sung and Kyomin Jung . . . . .                                                | 33422 |
| <i>CompassVerifier: A Unified and Robust Verifier for LLMs Evaluation and Outcome Reward</i>                                                                |       |
| Shudong Liu, Hongwei Liu, Junnan Liu, Linchen Xiao, Songyang Gao, Chengqi Lyu, Yuzhe Gu, Wenwei Zhang, Derek F. Wong, Songyang Zhang and Kai Chen . . . . . | 33455 |
| <i>A Knowledge-driven Adaptive Collaboration of LLMs for Enhancing Medical Decision-making</i>                                                              |       |
| Xiao Wu, Ting-Zhu Huang, Liang-Jian Deng, Yanyuan Qiao, Imran Razzak and Yutong Xie                                                                         | 33484 |
| <i>Castle: Causal Cascade Updates in Relational Databases with Large Language Models</i>                                                                    |       |
| Yongye Su, Yucheng Zhang, Zeru Shi, Bruno Ribeiro and Elisa Bertino . . . . .                                                                               | 33502 |



|                                                                                                                                                                                                                   |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Idiosyncratic Versus Normative Modeling of Atypical Speech Recognition: Dysarthric Case Studies</i>                                                                                                            |       |
| Vishnu Raja, Adithya V Ganesan, Anand Syamkumar, Ritwik Banerjee and H. Schwartz . .                                                                                                                              | 33515 |
| <i>NESTFUL: A Benchmark for Evaluating LLMs on Nested Sequences of API Calls</i>                                                                                                                                  |       |
| Kinjal Basu, Ibrahim Abdelaziz, Kiran Kate, Mayank Agarwal, Maxwell Crouse, Yara Rizk, Kelsey Bradford, Asim Munawar, Sadhana Kumaravel, Saurabh Goyal, Xin Wang, Luis A. Lastras and Pavan Kapanipathi . . . . . | 33527 |
| <i>Benchmarking and Mitigating MCQA Selection Bias of Large Vision-Language Models</i>                                                                                                                            |       |
| Md. Atabuzzaman, Ali Asgarov and Chris Thomas . . . . .                                                                                                                                                           | 33537 |
| <i>Can Large Language Models Unlock Novel Scientific Research Ideas?</i>                                                                                                                                          |       |
| Sandeep Kumar, Tirthankar Ghosal, Vinayak Goyal and Asif Ekbali . . . . .                                                                                                                                         | 33552 |
| <i>Word Salad Chopper: Reasoning Models Waste A Ton Of Decoding Budget On Useless Repetitions, Self-Knowingly</i>                                                                                                 |       |
| Wenya Xie, Shaochen Zhong, Hoang Anh Duy Le, Zhaozhuo Xu, Jianwen Xie and Zirui Liu                                                                                                                               | 33577 |
| <i>DIWALI - Diversity and Inclusivity aWare cuLture specific Items for India: Dataset and Assessment of LLMs for Cultural Text Adaptation in Indian Context</i>                                                   |       |
| Pramit Sahoo, Maharaj Brahma and Maunendra Sankar Desarkar . . . . .                                                                                                                                              | 33588 |
| <i>SYNC: A Synthetic Long-Context Understanding Benchmark for Controlled Comparisons of Model Capabilities</i>                                                                                                    |       |
| Shuyang Cao, Kaijian Zou and Lu Wang . . . . .                                                                                                                                                                    | 33616 |
| <i>OpenNER 1.0: Standardized Open-Access Named Entity Recognition Datasets in 50+ Languages</i>                                                                                                                   |       |
| Chester Palen-Michel, Maxwell Pickering, Maya Kruse, Jonne Sällevä and Constantine Lignos                                                                                                                         | 33638 |
| <i>Mondrian: A Framework for Logical Abstract (Re)Structuring</i>                                                                                                                                                 |       |
| Elizabeth Grace Orwig, Shinwoo Park, Hyundong Jin and Yo-Sub Han . . . . .                                                                                                                                        | 33664 |
| <i>Case-Based Decision-Theoretic Decoding with Quality Memories</i>                                                                                                                                               |       |
| Hiroyuki Deguchi and Masaaki Nagata . . . . .                                                                                                                                                                     | 33680 |
| <i>PRIME: Large Language Model Personalization with Cognitive Dual-Memory and Personalized Thought Process</i>                                                                                                    |       |
| Xinliang Frederick Zhang, Nicholas Beauchamp and Lu Wang . . . . .                                                                                                                                                | 33696 |
| <i>Mechanisms vs. Outcomes: Probing for Syntax Fails to Explain Performance on Targeted Syntactic Evaluations</i>                                                                                                 |       |
| Ananth Agarwal, Jasper Jian, Christopher D Manning and Shikhar Murty . . . . .                                                                                                                                    | 33726 |
| <i>Image Difference Captioning via Adversarial Preference Optimization</i>                                                                                                                                        |       |
| Zihan Huang, Junda Wu, Rohan Surana, Tong Yu, David Arbour, Ritwik Sinha and Julian McAuley . . . . .                                                                                                             | 33747 |
| <i>seqBench: A Tunable Benchmark to Quantify Sequential Reasoning Limits of LLMs</i>                                                                                                                              |       |
| Mohammad Ramezanali, Mo Vazifeh and Paolo Santi . . . . .                                                                                                                                                         | 33760 |
| <i>NormGenesis: Multicultural Dialogue Generation via Exemplar-Guided Social Norm Modeling and Violation Recovery</i>                                                                                             |       |
| Minki Hong, Jangho Choi and Jihie Kim . . . . .                                                                                                                                                                   | 33782 |

|                                                                                                                                                         |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>SATBench: Benchmarking LLMs’ Logical Reasoning via Automated Puzzle Generation from SAT Formulas</i>                                                 |       |
| Anjiang Wei, Yuheng Wu, Yingjia Wan, Tarun Suresh, Huanmi Tan, Zhanke Zhou, Sanmi Koyejo, Ke Wang and Alex Aiken . . . . .                              | 33821 |
| <i>Data Descriptions from Large Language Models with Influence Estimation</i>                                                                           |       |
| Chaeri Kim, Jaeyeon Bae and Taehwan Kim . . . . .                                                                                                       | 33839 |
| <i>EquiBench: Benchmarking Large Language Models’ Reasoning about Program Semantics via Equivalence Checking</i>                                        |       |
| Anjiang Wei, Jiannan Cao, Ran Li, Hongyu Chen, Yuhui Zhang, Ziheng Wang, Yuan Liu, Thiago S. F. X. Teixeira, Diyi Yang, Ke Wang and Alex Aiken. . . . . | 33857 |
| <i>MicroEdit: Neuron-level Knowledge Disentanglement and Localization in Lifelong Model Editing</i>                                                     |       |
| Shiqi Wang, Qi Wang, Runliang Niu, He Kong and Yi Chang . . . . .                                                                                       | 33871 |
| <i>Do Large Language Models Understand Word Senses?</i>                                                                                                 |       |
| Domenico Meconi, Simone Stirpe, Federico Martelli, Leonardo Lavallo and Roberto Navigli                                                                 | 33886 |
| <i>Diverse, not Short: A Length-Controlled Data Selection Strategy for Improving Response Diversity of Language Models</i>                              |       |
| Vijeta Deshpande, Debasmita Ghose, John D Patterson, Roger E. Beaty and Anna Rumshisky                                                                  | 33906 |
| <i>Uncovering the Bigger Picture: Comprehensive Event Understanding Via Diverse News Retrieval</i>                                                      |       |
| Yixuan Tang, Yuanyuan Shi, Yiqun Sun and Anthony Kum Hoe Tung . . . . .                                                                                 | 33928 |
| <i>Personalized LLM Decoding via Contrasting Personal Preference</i>                                                                                    |       |
| Hyungjune Bu, ChanJoo Jung, Minjae Kang and Jaehyung Kim . . . . .                                                                                      | 33947 |
| <i>The Missing Parts: Augmenting Fact Verification with Half Truth Detection</i>                                                                        |       |
| Yixuan Tang, Jincheng Wang and Anthony Kum Hoe Tung . . . . .                                                                                           | 33968 |
| <i>Toward Machine Translation Literacy: How Lay Users Perceive and Rely on Imperfect Translations</i>                                                   |       |
| Yimin Xiao, Yongle Zhang, Dayeon Ki, Calvin Bao, Marianna J. Martindale, Charlotte Vaughn, Ge Gao and Marine Carpuat . . . . .                          | 33986 |
| <i>Personalization up to a Point: Why Personalized Content Moderation Needs Boundaries, and How We Can Enforce Them</i>                                 |       |
| Emanuele Moscato, Tiancheng Hu, Matthias Orlikowski, Paul Röttger and Debora Nozza .                                                                    | 34004 |
| <i>MPCG: Multi-Round Persona-Conditioned Generation for Modeling the Evolution of Misinformation with LLMs</i>                                          |       |
| Chong Jun Rong Brian, Yixuan Tang and Anthony Kum Hoe Tung . . . . .                                                                                    | 34019 |
| <i>LiTeX: A Linguistic Taxonomy of Explanations for Understanding Within-Label Variation in Natural Language Inference</i>                              |       |
| Pingjun Hong, Beiduo Chen, Siyao Peng, Marie-Catherine de Marneffe and Barbara Plank                                                                    | 34054 |
| <i>LiteraryQA: Towards Effective Evaluation of Long-document Narrative QA</i>                                                                           |       |
| Tommaso Bonomo, Luca Gioffré and Roberto Navigli . . . . .                                                                                              | 34075 |
| <i>FillerSpeech: Towards Human-Like Text-to-Speech Synthesis with Filler Insertion and Filler Style Control</i>                                         |       |
| Seung-Bin Kim, Jun-Hyeok Cha, Hyung-Seok Oh, Heejin Choi and Seong-Whan Lee . . . .                                                                     | 34097 |



|                                                                                                                                                                                                                               |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Multi-LMentry: Can Multilingual LLMs Solve Elementary Tasks Across Languages?</i>                                                                                                                                          |       |
| Luca Moroni, Javier Aula-Blasco, Simone Conia, Irene Baucells, Naiara Perez, Silvia Paniagua Suárez, Anna Sallés, Malte Ostendorff, Júlia Falcão, Guijin Son, Aitor Gonzalez-Agirre, Roberto Navigli and Marta Villegas ..... | 34115 |
| <i>Lookahead Q-Cache: Achieving More Consistent KV Cache Eviction via Pseudo Query</i>                                                                                                                                        |       |
| Yixuan Wang, Shiyu Ji, Yijun Liu, Yuzhuang Xu, Yang Xu, Qingfu Zhu and Wanxiang Che                                                                                                                                           | 34147 |
| <i>PerspectiveMod: A Perspectivist Resource for Deliberative Moderation</i>                                                                                                                                                   |       |
| Eva Maria Vecchi, Neele Falk, Carlotta Quensel, Iman Jundi and Gabriella Lapesa .....                                                                                                                                         | 34164 |
| <i>LoCt-Instruct: An Automatic Pipeline for Constructing Datasets of Logical Continuous Instructions</i>                                                                                                                      |       |
| Hongyu Sun, Yusuke Sakai, Haruki Sakajo, Shintaro Ozaki, Kazuki Hayashi, Hidetaka Kamigaito and Taro Watanabe .....                                                                                                           | 34188 |
| <i>CodeSSM: Towards State Space Models for Code Understanding</i>                                                                                                                                                             |       |
| Shweta Verma, Abhinav Anand and Mira Mezini .....                                                                                                                                                                             | 34208 |
| <i>EduAdapt: A Question Answer Benchmark Dataset for Evaluating Grade-Level Adaptability in LLMs</i>                                                                                                                          |       |
| Numaan Naeem, Abdellah El Mekki and Muhammad Abdul-Mageed .....                                                                                                                                                               | 34225 |
| <i>xCoRe: Cross-context Coreference Resolution</i>                                                                                                                                                                            |       |
| Giuliano Martinelli, Bruno Gatti and Roberto Navigli .....                                                                                                                                                                    | 34253 |
| <i>Retrieval-Augmented Generation with Estimation of Source Reliability</i>                                                                                                                                                   |       |
| Jeongyeon Hwang, Junyoung Park, Hyejin Park, Dongwoo Kim, Sangdon Park and Jungseul Ok                                                                                                                                        | 34268 |
| <i>NitiBench: Benchmarking LLM Frameworks on Thai Legal Question Answering Capabilities</i>                                                                                                                                   |       |
| Pawitsapak Akarajaradwong, Pirat Pothavorn, Chompakorn Chaksangchaichot, Panuthep Tasawong, Thitiwat Nopparatbundit, Keerakiat Pratai and Sarana Nutanong .....                                                               | 34293 |
| <i>From Input Perception to Predictive Insight: Modeling Model Blind Spots Before They Become Errors</i>                                                                                                                      |       |
| Maggie Mi, Aline Villavicencio and Nafise Sadat Moosavi .....                                                                                                                                                                 | 34317 |
| <i>Wojood<sup>Relations</sup>: Arabic Relation Extraction Corpus and Modeling</i>                                                                                                                                             |       |
| Alaa Aljabari, Mohammed Khalilia and Mustafa Jarrar .....                                                                                                                                                                     | 34331 |
| <i>Conflicting Needles in a Haystack: How LLMs behave when faced with contradictory information</i>                                                                                                                           |       |
| Murathan Kurfali .....                                                                                                                                                                                                        | 34350 |
| <i>Towards Event Extraction with Massive Types: LLM-based Collaborative Annotation and Partitioning Extraction</i>                                                                                                            |       |
| Wenxuan Liu, Zixuan Li, Long Bai, Yuxin Zuo, Daozhu Xu, Xiaolong Jin, Jiafeng Guo and Xueqi Cheng .....                                                                                                                       | 34366 |
| <i>Liaozhai through the Looking-Glass: On Paratextual Explicitation of Culture-Bound Terms in Machine Translation</i>                                                                                                         |       |
| Sherrie Shen, Weixuan Wang and Alexandra Birch .....                                                                                                                                                                          | 34389 |
| <i>Concept-pedia: a Wide-coverage Semantically-annotated Multimodal Dataset</i>                                                                                                                                               |       |
| Karim Ghonim, Andrei Stefan Bejgu, Alberte Fernández-Castro and Roberto Navigli .....                                                                                                                                         | 34406 |
| <i>RAED: Retrieval-Augmented Entity Description Generation for Emerging Entity Linking and Disambiguation</i>                                                                                                                 |       |
| Karim Ghonim, Pere-Lluís Huguet Cabot, Riccardo Orlando and Roberto Navigli .....                                                                                                                                             | 34428 |

|                                                                                                                                                                                                                                            |       |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Personalized Language Models via Privacy-Preserving Evolutionary Model Merging</i><br>Kyuyoung Kim, Jinwoo Shin and Jaehyung Kim .....                                                                                                  | 34442 |
| <i>Aligning Text/Speech Representations from Multimodal Models with MEG Brain Activity During Listening</i><br>Padakanti Srijith, Khushbu Pahwa, Radhika Mamidi, Bapi Raju Surampudi, Manish Gupta and Subba Reddy Oota .....              | 34458 |
| <i>STARQA: A Question Answering Dataset for Complex Analytical Reasoning over Structured Databases</i><br>Mounica Maddela, Lingjue Xie, Daniel Preotiuc-Pietro and Mausam .....                                                            | 34476 |
| <i>Slim-SC: Thought Pruning for Efficient Scaling with Self-Consistency</i><br>Colin Hong, Xu Guo, Anand Chanaan Singh, Esha Choukse and Dmitrii Ustiugov .....                                                                            | 34489 |
| <i>Long Chain-of-Thought Fine-tuning via Understanding-to-Reasoning Transition</i><br>Chenxin An, Zhihui Xie, Xiaonan Li, Ming Zhong, Shansan Gong, Lei Li, Jun Zhang, Jingjing Xu and Lingpeng Kong .....                                 | 34507 |
| <i>Exploring Large Language Models for Detecting Mental Disorders</i><br>Gleb Kuzmin, Petr Strepotov, Maksim Stankevich, Natalia Chudova, Artem Shelmanov and Ivan Smirnov .....                                                           | 34524 |
| <i>Efficient Real-time Refinement of Language Model Text Generation</i><br>Joonho Ko, Jinheon Baek and Sung Ju Hwang .....                                                                                                                 | 34549 |
| <i>Reward-Weighted Sampling: Enhancing Non-Autoregressive Characteristics in Masked Diffusion LLMs</i><br>Daehoon Gwak, Minseo Jung, Junwoo Park, Minh Park, ChaeHun Park, Junha Hyung and Jaegul Choo .....                               | 34563 |
| <i>AI Argues Differently: Distinct Argumentative and Linguistic Patterns of LLMs in Persuasive Contexts</i><br>Esra Dönmez, Maximilian Maurer, Gabriella Lapesa and Agnieszka Falenska .....                                               | 34584 |
| <i>TounsiBench: Benchmarking Large Language Models for Tunisian Arabic</i><br>Souha Ben Hassine, Asma Arrak, Marouene Addhoum and Steven R Wilson .....                                                                                    | 34616 |
| <i>Moral Framing in Politics (MFIP): A new resource and models for moral framing</i><br>Ines Rehbein, Ines Reinig and Simone Paolo Ponzetto .....                                                                                          | 34632 |
| <i>ReDepress: A Cognitive Framework for Detecting Depression Relapse from Social Media</i><br>Aakash Kumar Agarwal, Saprativa Bhattacharjee, Mauli Rastogi, Jemima S. Jacob, Biplab Banerjee, Rashmi Gupta and Pushpak Bhattacharyya ..... | 34653 |
| <i>iKnow-audio: Integrating Knowledge Graphs with Audio-Language Models</i><br>Michel Olvera, Changhong Wang, Paraskevas Stamatiadis, Gaël Richard and Slim Essid ..                                                                       | 34672 |
| <i>EduVidQA: Generating and Evaluating Long-form Answers to Student Questions based on Lecture Videos</i><br>Sourjyadip Ray, Shubham Sharma, Somak Aditya and Pawan Goyal .....                                                            | 34690 |
| <i>The Illusion of Progress: Re-evaluating Hallucination Detection in LLMs</i><br>Denis Janiak, Jakub Binkowski, Albert Sawczyn, Bogdan Gabrys, Ravid Shwartz-Ziv and Tomasz Jan Kajdanowicz .....                                         | 34717 |
| <i>Turning Logic Against Itself: Probing Model Defenses Through Contrastive Questions</i><br>Rachneet Singh Sachdeva, Rima Hazra and Iryna Gurevych .....                                                                                  | 34735 |

|                                                                                                                                                                                                   |       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>CHURRO: Making History Readable with an Open-Weight Large Vision-Language Model for High-Accuracy, Low-Cost Historical Text Recognition</i>                                                    |       |
| Sina Semnani, Han Zhang, Xinyan He, Merve Tekgurler and Monica Lam                                                                                                                                | 34766 |
| <i>Towards Author-informed NLP: Mind the Social Bias</i>                                                                                                                                          |       |
| Inbar Pendzel and Einat Minkov                                                                                                                                                                    | 34814 |
| <i>Detecting Corpus-Level Knowledge Inconsistencies in Wikipedia with Large Language Models</i>                                                                                                   |       |
| Sina Semnani, Jirayu Burapachee, Arpandee Khatua, Thanawan Atchariyachanvanit, Zheng Wang and Monica Lam                                                                                          | 34828 |
| <i>Leveraging Multilingual Training for Authorship Representation: Enhancing Generalization across Languages and Domains</i>                                                                      |       |
| Junghwan Kim, Haotian Zhang and David Jurgens                                                                                                                                                     | 34856 |
| <i>DrFrattn: Directly Learn Adaptive Policy from Attention for Simultaneous Machine Translation</i>                                                                                               |       |
| Libo Zhao, Jing Li and Ziqian Zeng                                                                                                                                                                | 34882 |
| <i>The Sound of Syntax: Finetuning and Comprehensive Evaluation of Language Models for Speech Pathology</i>                                                                                       |       |
| Fagun Patel, Duc Quang Nguyen, Sang T. Truong, Jody Vaynshtok, Sanmi Koyejo and Nick Haber                                                                                                        | 34896 |
| <i>NormXLogit: The Head-on-Top Never Lies</i>                                                                                                                                                     |       |
| Sina Abbasi, Mohammad Reza Modarres and Mohammad Taher Pilehvar                                                                                                                                   | 34915 |
| <i>Doc2Chart: Intent-Driven Zero-Shot Chart Generation from Documents</i>                                                                                                                         |       |
| Akriti Jain, Pritika Ramu, Aparna Garimella and Apoorv Saxena                                                                                                                                     | 34937 |
| <i>Breaking Agents: Compromising Autonomous LLM Agents Through Malfunction Amplification</i>                                                                                                      |       |
| Boyang Zhang, Yicong Tan, Yun Shen, Ahmed Salem, Michael Backes, Savvas Zannettou and Yang Zhang                                                                                                  | 34953 |
| <i>ForEST: Frame of Reference Evaluation in Spatial Reasoning Tasks</i>                                                                                                                           |       |
| Tanawan Premisri and Parisa Kordjamshidi                                                                                                                                                          | 34966 |
| <i>Multilinguality Does not Make Sense: Investigating Factors Behind Zero-Shot Cross-Lingual Transfer in Sense-Aware Tasks</i>                                                                    |       |
| Roksana Goworek and Haim Dubossarsky                                                                                                                                                              | 34993 |
| <i>Translating Domain-Specific Terminology in Typologically-Diverse Languages: A Study in Tax and Financial Education</i>                                                                         |       |
| Arturo Oncevay, Elena Kochkina, Keshav Ramani, Toyin Aguda, Simerjot Kaur and Charese Smiley                                                                                                      | 35019 |
| <i>Train It and Forget It: Merge Lists are Unnecessary for BPE Inference in Language Models</i>                                                                                                   |       |
| Tomohiro Sawada and Kartik Goyal                                                                                                                                                                  | 35034 |
| <i>Spectral Scaling Laws in Language Models: <i>emph</i>How Effectively Do Feed-Forward Networks Use Their Latent Space?</i>                                                                      |       |
| Nandan Kumar Jha and Brandon Reagan                                                                                                                                                               | 35048 |
| <i>TLUE: A Tibetan Language Understanding Evaluation Benchmark</i>                                                                                                                                |       |
| Fan Gao, Cheng Huang, Yutong Liu, Nyima Tashi, Xiangxiang Wang, Thupten Tsering, Ban Ma-bao, Renzeng Duo jie, Gadeng Luosang, Rinchen Dongrub, Dorje Tashi, Xiao Feng CD, Yongbin Yu and Hao Wang | 35060 |

|                                                                                                                                                                                                                                                                                                                                   |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| <i>Retrieving Support to Rank Answers in Open-Domain Question Answering</i><br>Zeyu Zhang, Alessandro Moschitti and Thuy Vu .....                                                                                                                                                                                                 | 35087 |
| <i>Trojsten Benchmark: Evaluating LLM Problem-Solving in Slovak STEM Competition Problems</i><br>Adam Zahradník and Marek Suppa .....                                                                                                                                                                                             | 35095 |
| <i>BRSpeech-DF: A Deep Fake Synthetic Speech Dataset for Portuguese Zero-Shot TTS</i><br>Alexandre Costa Ferro Filho, Rafaello Virgilli, Lucas Alcantara Souza, F S de Oliveira, Marcelo Henrique Lopes Ferreira, Daniel Tunnermann, Gustavo Dos Reis Oliveira, Anderson Da Silva Soares and Arlindo Rodrigues Galvão Filho ..... | 35111 |
| <i>A Simple Yet Effective Method for Non-Refusing Context Relevant Fine-grained Safety Steering in LLMs</i><br>Shaona Ghosh, Amrita Bhattacharjee, Yftah Ziser and Christopher Parisien .....                                                                                                                                     | 35117 |
| <i>Statistical and Neural Methods for Hawaiian Orthography Modernization</i><br>Jaden Kapali, Keaton Williamson and Winston Wu .....                                                                                                                                                                                              | 35138 |
| <i>so much depends / upon / a whitespace: Why Whitespace Matters for Poets and LLMs</i><br>Sriharsh Bhyravajula, Melanie Walsh, Anna Preus and Maria Antoniak .....                                                                                                                                                               | 35145 |
| <i>Certified Mitigation of Worst-Case LLM Copyright Infringement</i><br>Jingyu Zhang, Jiacan Yu, Marc Marone, Benjamin Van Durme and Daniel Khashabi .....                                                                                                                                                                        | 35163 |
| <i>Quantifying Logical Consistency in Transformers via Query-Key Alignment</i><br>Eduard Tulchinskii, Laida Kushnareva, Anastasia Voznyuk, Andrei Andriiainen, Irina Piontkovskaya, Evgeny Burnaev and Serguei Barannikov .....                                                                                                   | 35185 |
| <i>SimulatorArena: Are User Simulators Reliable Proxies for Multi-Turn Evaluation of AI Assistants?</i><br>Yao Dou, Michel Galley, Baolin Peng, Chris Kedzie, Weixin Cai, Alan Ritter, Chris Quirk, Wei Xu and Jianfeng Gao .....                                                                                                 | 35201 |
| <i>CourtReasoner: Can LLM Agents Reason Like Judges?</i><br>Sophia Simeng Han, Yoshiki Takashima, Shannon Zejiang Shen, Chen Liu, Yixin Liu, Roque K. Thuo, Sonia Knowlton, Ruzica Piskac, Scott J Shapiro and Arman Cohan .....                                                                                                  | 35280 |
| <i>Not Your Typical Government Tipline: LLM-Assisted Routing of Environmental Protection Agency Citizen Tips</i><br>Sharanya Majumder, Zehua Li, Derek Ouyang, Kit T Rodolfa, Elena Eneva, Julian Nyarko and Daniel E. Ho .....                                                                                                   | 35296 |
| <i>Retracing the Past: LLMs Emit Training Data When They Get Lost</i><br>Myeongseob Ko, Nikhil Reddy Billa, Adam Nguyen, Charles Fleming, Ming Jin and Ruoxi Jia .....                                                                                                                                                            | 35305 |
| <i>Layer-wise Minimal Pair Probing Reveals Contextual Grammatical-Conceptual Hierarchy in Speech Representations</i><br>Linyang He, Qiaolin Wang, Xilin Jiang and Nima Mesgarani .....                                                                                                                                            | 35327 |
| <i>Current Semantic-change Quantification Methods Struggle with Semantic Change Discovery in the Wild</i><br>Khonzoda Umarova, Lillian Lee and Laerdon Kim .....                                                                                                                                                                  | 35343 |
| <i>Evaluating Large Language Models for Detecting Antisemitism</i><br>Jay Patel, Hrudayangam Mehta and Jeremy Blackburn .....                                                                                                                                                                                                     | 35357 |
| <i>D-RAG: Differentiable Retrieval-Augmented Generation for Knowledge Graph Question Answering</i><br>Guangze Gao, Zixuan Li, Chunfeng Yuan, Jiawei Li, Wu Jianzhuo, Yuehao Zhang, Xiaolong Jin, Bing Li and Weiming Hu .....                                                                                                     | 35387 |

*Towards Robust Mathematical Reasoning*

Thang Luong, Dawsen Hwang, Hoang H Nguyen, Golnaz Ghiasi, Yuri Chervonyi, Insuk Seo, Jun-su Kim, Garrett Bingham, Jonathan Lee, Swaroop Mishra, Alex Zhai, Huiyi Hu, Henryk Michalewski, Jimin Kim, Jeonghyun Ahn, Junhwi Bae, Xingyou Song, Trieu Hoang Trinh, Quoc V Le and Junehyuk Jung ..... 35407

*Table-LLM-Specialist: Language Model Specialists for Tables using Iterative Fine-tuning*

Junjie Xing, Yeye He, Mengyu Zhou, Haoyu Dong, Shi Han, Dongmei Zhang and Surajit Chaudhuri ..... 35432

*Introducing Spotlight: A Novel Approach for Generating Captivating Key Information from Documents*

Ankan Mullick, Sombit Bose, Rounak Saha, Ayan Kumar Bhowmick, Aditya Vempaty, Prasenjit Dey, Ravi Kokku, Pawan Goyal and Niloy Ganguly ..... 35450

*Argument Summarization and its Evaluation in the Era of Large Language Models*

Moritz Altemeyer, Steffen Eger, Johannes Daxenberger, Yanran Chen, Tim Altendorf, Philipp Cimiano and Benjamin Schiller ..... 35479

*Computational Analysis of Conversation Dynamics through Participant Responsivity*

Margaret Hughes, Brandon Roy, Elinor Poole-Dayana, Deb Roy and Jad Kabbara ..... 35501

*AMQ: Enabling AutoML for Mixed-precision Weight-Only Quantization of Large Language Models*

Sangjun Lee, Seung-taek Woo, Jun-gyu Jin, Changhun Lee and Eunhyeok Park ..... 35521

*Beyond Averages: Learning with Annotator Disagreement in STS*

Alejandro Benito-Santos and Adrian Ghajari ..... 35540

*Dipper: Diversity in Prompts for Producing Large Language Model Ensembles in Reasoning Tasks*

Wenyang Hu, Gregory Kang Ruey Lau, Liu Diwen, Chen Jizhuo, See-Kiong Ng and Bryan Kian Hsiang Low ..... 35547

*Constrained Non-negative Matrix Factorization for Guided Topic Modeling of Minority Topics*

Seyedeh Fatemeh Ebrahimi and Jaakko Peltonen ..... 35562

*Which Word Orders Facilitate Length Generalization in LMs? An Investigation with GCG-Based Artificial Languages*

Nadine El-Naggar, Tatsuki Kuribayashi and Ted Briscoe ..... 35588

*Training compute-optimal transformer encoder models*

Megi Dervishi, Alexandre Allauzen, Gabriel Synnaeve and Yann LeCun ..... 35603

*Mind the Blind Spots: A Focus-Level Evaluation Framework for LLM Reviews*

Hyungyu Shin, Jingyu Tang, Yoonjoo Lee, Nayoung Kim, Hyunseung Lim, Ji Yong Cho, Hwajung Hong, Moontae Lee and Juho Kim ..... 35619

*Seeing Through Words, Speaking Through Pixels: Deep Representational Alignment Between Vision and Language Models*

Zoe Wanying He, Sean Trott and Meenakshi Khosla ..... 35646

*Unconditional Truthfulness: Learning Unconditional Uncertainty of Large Language Models*

Artem Vazhentsev, Ekaterina Fadeeva, Rui Xing, Gleb Kuzmin, Ivan Lazichny, Alexander Panchenko, Preslav Nakov, Timothy Baldwin, Maxim Panov and Artem Shelmanov ..... 35662

*Chinese Toxic Language Mitigation via Sentiment Polarity Consistent Rewrites*

Xintong Wang, Yixiao Liu, Jingheng Pan, Liang Ding, Longyue Wang and Chris Biemann ..... 35684

*A Head to Predict and a Head to Question: Pre-trained Uncertainty Quantification Heads for Hallucination Detection in LLM Outputs*

Artem Shelmanov, Ekaterina Fadeeva, Akim Tsvigun, Ivan Tsvigun, Zhuohan Xie, Igor Kiselev, Nico Daheim, Caiqi Zhang, Artem Vazhentsev, Mrinmaya Sachan, Preslav Nakov and Timothy Baldwin  
35701