



COLING 2025 • Abu Dhabi

31st International Conference on Computational Linguistics

JANUARY 19-24

Virtual Poster Sessions hosted in Gather
Monday, January 27th and Tuesday, January 28th

Contents

| | |
|--|-----------|
| Table of Contents | i |
| 1 Conference Information | 1 |
| Message from the General Chairs | 1 |
| Message from the Program Chairs | 3 |
| Message from the Local Chairs | 4 |
| Organizing Committee | 6 |
| Senior Program Committee | 8 |
| Conference Organizers & Vendors | 11 |
| 2 Diversity, Inclusion, and Anti-Harassment Policy | 13 |
| 3 Meal Info | 15 |
| 4 Welcome Reception | 17 |
| 5 Social Events – Offsite Excursion | 19 |
| Social Excursion Option 1: Grand Mosque and Presidential Palace Afternoon Tour | 21 |
| Social Excursion Option 2: Abrahamic House and Louvre Abu Dhabi Afternoon Tour | 22 |
| Social Excursion Options 3: Sand Dunes & 4x4 Desert Safari | 24 |
| 6 Keynotes | 27 |
| 7 Birds-of-a-Feather Group Discussions | 31 |
| 8 Conference Overview | 33 |
| Saturday, January 18, 2025 | 33 |
| Sunday, January 19, 2025 | 33 |
| Monday, January 20, 2025 | 33 |
| Tuesday, January 21, 2025 | 34 |

| | |
|--|-----------|
| Wednesday, January 22, 2025 | 35 |
| Thursday, January 23, 2025 | 36 |
| Friday, January 24, 2025 | 37 |
| 9 Tutorials Overview | 39 |
| Tutorials Overview: Sunday, January 19, 2025 | 39 |
| Tutorial Overviews: Monday, January 20, 2025 | 40 |
| 10 Tutorials Details | 41 |
| Sunday, January 19, 2025 | 41 |
| Monday, January 20, 2025 | 43 |
| 11 Workshops Overview | 45 |
| Workshops Overview: Sunday, January 19, 2025 | 45 |
| Workshops Overview: Sunday-Monday, January 19-20, 2025 | 46 |
| Workshops Overview: Sunday-Monday, January 20, 2025 | 46 |
| 12 Workshop Details | 49 |
| Sunday, January 19, 2025 | 49 |
| Sunday-Monday, January 19-20, 2025 | 52 |
| Monday, January 20, 2025 | 52 |
| 13 Local Guide | 59 |
| Conference Venue | 59 |
| Abu Dhabi | 60 |
| Abu Dhabi Landmarks | 62 |
| Important Information | 63 |
| Dubai | 65 |
| Sharjah | 66 |
| 14 Venue Map | 67 |
| Author Index | 69 |
| Sponsorship | 69 |

Conference Information

Message from the General Chairs

On behalf of the International Committee on Computational Linguistics (ICCL), we are delighted to welcome you to the 31st International Conference on Computational Linguistics: COLING 2025, held in Abu Dhabi, United Arab Emirates.

Unlike the previous 30 editions of COLING, which were held every two years, this COLING takes place only eight months after the preceding COLING conference in Turin, Italy, which was (exceptionally) joint with the LREC conference. What might look like a not very clever move (to say the least) – especially in view of the density of the major conferences in our field in summer and autumn – was motivated by the goal to disentangle LREC and COLING, the two major biannual conferences, which have, despite all their differences, a lot in common and therefore also share a significant percentage of attendees. From now on, COLING will be held in odd years (as had already been the case from 1965 to 1973), while LREC will continue to be held in even years. In addition to this change, COLING 2025 introduces another novelty: the main conference is in-person only, but, in order to facilitate the participation of researchers who cannot travel to Abu Dhabi for whatever reason, we have introduced a separate, entirely virtual component of the conference. The virtual poster session will take place the week after the in-person part, and will enable all authors to present their papers.

Our colleagues from the Mohamed bin Zayed University of Artificial Intelligence in Abu Dhabi generously offered to host COLING 2025, and what month is better to hold a conference in the United Arab Emirates than January! And here we are! Obviously, the short organization time has been a challenge for the entire Organizing Committee. The Program Chairs Hend Al-Khalifa, Marianna Apidianaki, Barbara Di Eugenio, and Steven Schockaert, have managed to organize the submission, reviewing, selection, and scheduling of a large number of papers in an exceptionally short time frame. Our Local Committee Chairs Preslav Nakov, Nizar Habash, Yova Kementchedjheva, Monojit Choudhury, Ahmed Dabbagh, Ekaterina Kochmar, Alham Fikri Aji, and Teresa Lynn have paid attention to all other organizational details (venue, accommodations, visas, budget). Especially Preslav was in demand in all matters, and in all of our numerous meetings. Jenn Rachford (R & R Planning) took over a substantial part of the organization of the conference and guided us with her professionalism, experience and wisdom.

However, we must not forget the hard work by the other members of our Organizing Committee:

Workshop Chairs: Katsuhito Sudoh and Mo El-Haj; Tutorial Chairs: Djámé Seddah and Shaonan Wang; Demonstration Chairs: Brodie Mather and Mark Dras; Industrial Track Chairs: Kareem Darwish and Apoorv Agarwal; Ethics Chairs: Zeerak Talat and Su Lin Blodgett; Publicity & Social Media Chairs: Sakhar AlKhereyf and Baotian Hu; Website Chairs: Hasan Iqbal and Zain Muhammad Mujahid; Publication Chairs: Emily Allaway and Guillermo Cámbara; Volunteer Chair: Artem Shelmanov; Virtual Infrastructure Chairs: Tuhin Chakrabarty and Alexander Shvets; Birds-of-a-Feather Chairs: Hawau Olamide Toyin and Mukund Choudhary; and Best Paper Committee Chairs: Kevin Duh and Aline Villavicencio. All of them had to put in an extra work shift in order to be able to cope with the very demanding time constraints. We would like to take the opportunity to thank all of them! We could not have wished for a better or more dedicated team to work alongside. We would also like to thank our generous sponsors: The Abu Dhabi Convention and Exhibition Center, The Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), the Technology Innovation Institute, New York University Abu Dhabi, Inception, Baidu, Google, Toloka AI, and Applied AI.

The result of this tremendous joint effort is an exciting high-quality program of 853 papers at the main conference, in addition to 22 workshops, 9 tutorials, and a winter school. In the in-person part of the conference program, we maintain the spirit of COLING and offer a number of different excursions that, as we are sure, will stimulate the social interaction between the attendees and allow them to immerse themselves into the cultural attractions of the United Arab Emirates.

Last but not least, we would like to thank each and every one of you for making COLING 2025 such a success by submitting high quality papers and demos, organizing workshops and tutorials on hot topics in our field, serving as Senior Area Chairs, Area Chairs, and Reviewers or Invited Speakers, and by joining us in Abu Dhabi.

Enjoy COLING 2025 in Abu Dhabi!

Owen Rambow and Leo Wanner
General Chairs of COLING 2025

Message from the Program Chairs

Welcome to the 31st International Conference on Computational Linguistics!

COLING 2025 is an in-person event which will take place in the beautiful historical city of Abu Dhabi. The conference includes a virtual 2-day event which will take place online shortly after the in-person conference.

COLING 2025 received 2,530 direct submissions and 193 papers committed via the ACL Rolling Review (ARR) route. Of the direct submissions, 2,335 were reviewed: 1,922 long papers, of which 605 were accepted (31.5%), and 413 short papers, of which 85 were accepted (20.6%). From the papers that were committed via ARR, we accepted 58 long papers, out of a total of 167 (34.7%), and 12 short papers, out of a total of 26 (46.2%). Across all categories, we have accepted a total of 760 papers (30.1%). We are gratified to note the great interest and enthusiasm that researchers in our community have shown for COLING 2025. We would like to thank all authors who submitted their work to the conference, and congratulate those whose papers were accepted for presentation.

We would like to thank the amazing pool of 2,418 reviewers, who made the conference possible and devoted their time to the careful assessment of the papers assigned to them, and in particular the emergency reviewers who stepped in at the last moment. Much of the heavy lifting in the review process fell on the 62 Senior Area Chairs and the 274 Area Chairs; we are extremely grateful to them for their efforts and time. We also want to thank the OpenReview team for their help in handling papers submitted to COLING 2025 via ARR. We would like to express our gratitude to our Best Paper Awards Committee Chairs, Kevin Duh (Johns Hopkins University, USA) and Aline Villavicencio (University of Exeter and University of Sheffield (UK), Federal University of Rio Grande do Norte (Brazil)), and to the Awards Committee members.

The COLING 2025 program features three exciting keynote speakers: Katrin Erk (University of Texas at Austin, USA), Emmanuel Dupoux (EHESS, META, France), and Partha Talukdar (Google DeepMind and IISc Bangalore, India).

To everyone mentioned above, we again express our deepest thanks for helping in so many different ways to make COLING 2025 a memorable event. We hope that you will enjoy the conference!

Marianna Apidianaki
Hend Al-Khalifa
Barbara Di Eugenio
Steven Schockaert
COLING 2025 Program Committee Co-Chairs

Message from the Local Chairs

Dear COLING 2025 Participants,

We are happy to welcome you to the 31st International Conference on Computational Linguistics (COLING 2025), hosted in vibrant Abu Dhabi, the heart of Artificial Intelligence (AI), innovation, and culture in the United Arab Emirates.

COLING 2025 brings together the brightest minds in AI and Computational Linguistics, creating an unparalleled opportunity for dialogue and innovation. As language technology continues to redefine the way humans interact with machines, we are looking forward to exploring how these advancements can be leveraged to solve regional and global challenges—from healthcare and education to business and beyond.

Abu Dhabi offers a unique blend of tradition and modernity. From the awe-inspiring architecture of the Sheikh Zayed Grand Mosque to the cutting-edge developments on Saadiyat Island, the iconic Louvre Abu Dhabi, and the bustling cultural hub of Qasr Al Hosn, the city reflects the values of connectivity, learning, and exploration—values that resonate deeply with our scientific community.

As local chairs, our aim has been to create an environment where attendees can engage in meaningful discussions, form new collaborations, and experience the warmth of Emirati hospitality. We are proud to showcase our region’s dedication to technology and language innovation through the conference program, which features distinguished keynote speakers, dynamic workshops, and opportunities to explore diverse linguistic and cultural insights.

We encourage you to take advantage of this exceptional location, not only by participating in the academic program but also by immersing yourself in the cultural richness of Abu Dhabi. Whether you visit the Louvre Abu Dhabi, stroll along the Corniche, or sample traditional Arabic cuisine, we hope you leave with memories that extend far beyond the conference halls.

With this in mind, we have prepared excursions that showcase the best of Abu Dhabi’s culture, history, and adventure: (i) for history enthusiasts, a visit to the Abrahamic House & Louvre Abu Dhabi, (ii) for architecture and culture lovers, the Sheikh Zayed Grand Mosque & Qasr Al Watan (Presidential Palace), and (iii) for adventure seekers, dune bashing, sand surfing, and camel riding. All three excursions conclude with a traditional Arabic dinner in the desert.

The logo for COLING 2025, designed by Professor Nizar Habash from NYUAD & MBZUAI, showcases the architectural marvels of Abu Dhabi, UAE. It features, from left to right, seven iconic buildings, including the Sheikh Zayed Grand Mosque, Capital Gate, aka leaning tower of Abu Dhabi (near our conference location), Etihad Towers, Burj Mohammed bin Rashid, Al Bahr Twin Towers, and Aldar Headquarters Building.

As the local organizer of COLING 2025, Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), the world’s first specialized research-based AI university, we are glad to see for the first time this prestigious conference to be held in the MENA region. This is a notable milestone for the region, following the successful 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP), which was organized by NYU Abu Dhabi (NYUAD) in partnership with MBZUAI.

COLING 2025 is a dynamic platform for showcasing Abu Dhabi’s growing leadership in AI, especially in natural language processing. Abu Dhabi is the home of MBZUAI’s Institute of Foundation Models, which has developed: Jais, in collaboration with G42 and Cerebras, the world’s leading open Arabic-centric foundation and instruction-tuned large language model; K2 a 65B model optimized for reasoning;

Nanda, the world’s best open Hindi model; and LLM360, the first truly open LLM. It is also the home of TII, which has developed the Falcon family of LLMs.

Being in the MENA region, there is a special focus of COLING 2025 on local languages. In addition to the main conference, the tutorials, and the workshops, we further host the First Arabic Natural Language Processing (NLP) Winter School, a two-day event (sponsored by NYUAD, with the first day hosted in MBZUAI). The Winter School is designed to advance knowledge and skills in Arabic NLP research and development and to foster interdisciplinary exchange and networking among students, academics, and industry professionals in the field.

In addition to leveraging Abu Dhabi’s strong MICE (meetings, incentives, conferences, and exhibitions) infrastructure and capabilities, the conference is supported by the Abu Dhabi Convention and Exhibition Bureau (ADCEB), part of the Department of Tourism and Culture – Abu Dhabi (DCT Abu Dhabi).

In closing, we hope that you will enjoy your stay in Abu Dhabi, exploring its rich culture, taking advantage of the many opportunities for fun, all while getting the most out of the extensive technical program of COLING 2025.

Welcome to COLING 2025 and welcome to the vibrant city of Abu Dhabi!

Hayyakum fi Darkum ‘Welcome to our shared home’.

Preslav Nakov, Nizar Habash, Yova Kementchedjhieva, Monojit Choudhury, Ahmed Dabbagh, Ekaterina Kochmar, Alham Fikri Aji, and Teresa Lynn
Local Chairs, COLING 2025

Organizing Committee

General Chairs

Owen Rambow, Stony Brook University, USA
Leo Wanner, ICREA and Universitat Pompeu Fabra, Spain

Program Chairs

Marianna Apidianaki, University of Pennsylvania, USA
Hend Al-Khalifa, King Saud University, Saudi Arabia
Barbara Di Eugenio, University of Illinois Chicago, USA
Steven Schockaert, Cardiff University, UK

Local Organization Chairs

Preslav Nakov, Mohamed bin Zayed University of Artificial Intelligence, UAE
Alham Fikri Aji, Mohamed bin Zayed University of Artificial Intelligence, UAE
Yova Kementchedjheva, Mohamed bin Zayed University of Artificial Intelligence, UAE
Ekaterina Kochmar, Mohamed bin Zayed University of Artificial Intelligence, UAE
Teresa Lynn, Mohamed bin Zayed University of Artificial Intelligence, UAE
Monojit Choudhury, Mohamed bin Zayed University of Artificial Intelligence, UAE
Nizar Habash, New York University Abu Dhabi, UAE, and Mohamed bin Zayed University of Artificial Intelligence, UAE
Ahmed Dabbagh, Mohamed bin Zayed University of Artificial Intelligence, UAE

Workshops Chairs

Katsuhito Sudoh, Nara Institute of Science and Technology, Japan
Mo El-Haj, Lancaster University, UK

Tutorial Chairs

Djamé Seddah, Institut National de Recherche en Informatique et en Automatique, France
Shaonan Wang, Institute of Automation, Chinese Academy of Sciences, China

Demonstrations Chairs

Brodie Mather, The Institute of Human & Machine Cognition, USA
Mark Dras, Macquarie University, Australia

Industrial Track Chairs

Kareem Darwish, aiXplain, USA
Apoorv Agarwal, Columbia University, USA

Sponsorship Chairs

Monojit Choudhury, Mohamed bin Zayed University of Artificial Intelligence, UAE
Ahmed Dabbagh, Mohamed bin Zayed University of Artificial Intelligence, UAE

Ethics Chairs

Zeerak Talat, Mohamed bin Zayed University of Artificial Intelligence, UAE
Su Lin Blodgett, Microsoft Research

Publicity Chairs

Sakhar AlKhereyf, King Abdulaziz City for Science and Technology, Saudi Arabia
Baotian Hu, Harbin Institute of Technology, China

Website Chairs

Hasan Iqbal, Mohamed bin Zayed University of Artificial Intelligence, UAE
Zain Muhammad Mujahid, Mohamed bin Zayed University of Artificial Intelligence, UAE

Visa Chairs

Preslav Nakov, Mohamed bin Zayed University of Artificial Intelligence, UAE

Publications Chairs

Emily Allaway, University of Edinburgh, UK
Guillermo Cámara, Amazon

Volunteer Chairs

Artem Shelmanov, Mohamed bin Zayed University of Artificial Intelligence, UAE

Virtual Main Conference Chairs

Alexander Shvets, Barcelona Supercomputing Center, Spain
Tuhin Chakrabarty, Stony Brook University, USA

Best Paper Committee Chairs

Kevin Duh, Johns Hopkins University, USA
Aline Villavicencio, University of Sheffield, UK

Birds-of-a-Feather (BoF) Chairs

Hawau Olamide Toyin, Mohamed bin Zayed University of Artificial Intelligence, UAE
Mukund Choudhary, Mohamed bin Zayed University of Artificial Intelligence, UAE

Handbook Chair

Salam Khalifa, Stony Brook University, USA

Senior Program Committee

Linguistic Theories, Cognitive Modeling and Psycholinguistics

Abdellah Fourtassi, Aix-Marseille Université
Jordan Kodner, Stony Brook University
David Mortensen, Carnegie Mellon University (CMU)

Dialogue and Conversational Interaction

Chloé Clavel, INRIA
Wenjie Li, The Hong Kong Polytechnic University
Kallirroi Georgila, University of Southern California

Discourse and Pragmatics

Sharid Loáiciga, University of Gothenburg
Philippe Muller, University of Toulouse
Sebastian Schuster, University College London

Information Retrieval and Text Mining

Goran Glavaš, University of Würzburg
Wai Lam, Chinese University of Hong Kong
Pierre Zweigenbaum, University Paris-Saclay, LISN, CNRS

Ethics, Bias, and Fairness

Maria Antoniak, Allen Institute for AI
Danushka Bollegalla, University of Liverpool

Reasoning, Question Answering, and Sentence-level Semantics

Marie-Catherine de Marneffe, Université Catholique de Louvain
Sanda Harabagiu, University of Texas at Dallas
Niranjan Balasubramanian, Stony Brook University

Information Extraction

Naoaki Okazaki, Tokyo Institute of Technology
Qifan Wang, META AI

Language Modeling

Xiaodong Liu, Microsoft Research
Lun-Wei Ku, Academia Sinica
Soujanya Poria, University of Singapore

Language Resources and Evaluation

Serge Sharoff, University of Leeds

Valerio Basile, University of Turin
Nafise Sadat Moosavi, University of Sheffield

Lexical Semantics

Aina Garí Soler, Télécom Paris
Andrey Kutuzov, University of Oslo
Sebastian Padó, University of Stuttgart

Low-resourced and Less Studied Languages

Cristina España i Bonet, Universitat Politècnica de Catalunya (UPC)
Yves Scherrer, University of Oslo
Zoey Liu, University of Florida

Machine Learning for CL/NLP

João Sedoc, New York University (NYU)
Bing Liu, University of Illinois Chicago (UIC)
Christos Christodoulopoulos, Amazon

Multilinguality and Machine Translation

Benjamin Marie, The Kaitechup
Maja Popović, ADAPT Centre, Dublin City University

Phonology, Morphology, and Word Segmentation

Kemal Oflazer, Carnegie Mellon University
Eleanor Chodroff, University of Zurich
Ryan Cotterell, ETH Zürich

Syntax, Tagging, Chunking and Parsing

Yue Zhang, Westlake University
Miryam de Lhoneux, KU Leuven
Natalie Schluter, Apple and Technical University of Denmark

Multimodal and Grounded Language Acquisition, HRI

Casey Kennington, Boise State University
Malihe Alikhani, Northeastern University

Natural Language Generation, Summarization and Simplification

Kathy McKeown, Columbia University
Xiaojun Wan, Peking University
Fei Liu, Emory University

Offensive and Non-Inclusive Language Detection and Analysis

Sara Tonelli, Fondazione Bruno Kessler
Eduardo Blanco, University of Arizona

Sentiment Analysis, Opinion and Argument Mining

Elena Cabrio, Université Côte d'Azur
Manfred Stede, Potsdam University
Simone Paolo Ponzetto, Universität Mannheim

Speech Recognition and Synthesis and Spoken Language Understanding

Rita Singh, Carnegie Mellon University (CMU)
Frédéric Bechet, Aix-Marseille Université
Sakriani Sakti, Nara Institute of Science and Technology

Interpretability and Explainability

Nadir Durrani, Qatar Computing Research Institute
Carolin Lawrence, NEC Laboratories Europe

NLP and LLM Applications

Dina Demner Fushman, National Library of Medicine
Vincent Ng, University of Texas at Dallas
Joel Tetrault, Dataminr
Sophia Ananiadou, University of Manchester
Hai Zhao, Shanghai Jiao Tong University

Conference Organizers & Vendors

I would like to extend my sincere gratitude to Professors Timothy Baldwin, Preslav Nakov and the Local Committee for their invaluable assistance in the submission and selection process for our COLING 2025 venue. I also appreciate the support of the ICCL for their efforts in selecting MBZUAI as our conference Host and location.

Additionally, I would like to acknowledge our partners for COLING 2025:

- Onsite Team - Megan Haddad, Event Assistant; Lina Staggs, Support Staff
- Lee Hartman & Sons - Jon Dorsey, AV Director; Trevor Laffoon, AV; Brandy Dorsey, Support Staff
- ADNEC: Thembi Kuchena & Jyothi Joseph
- Emirates Tours & Safari LLC: Moheb Nazari
- Capital 360 Event Experience - Merlene Gomes, Maryam Al Baulushi, Maher Yahya
- Aloft: Kavita Tokekar, Poorva Kulkarni
- Andaz: Dimple Kalita

Thank you to the entire organizing and program committees for your hard work, dedication, and countless hours of effort sometimes at the expense of sleep & holidays making this conference a success.

Sincerely,
Jennifer Rachford
COLING 2025 Director of Events
R&R Planning

Diversity, Inclusion, and Anti-Harassment Policy

Diversity & Inclusion, both in terms of approaches and participation, are an integral part of who COLING is as a community. To find out more about the unique history of the COLING series of conferences visit: <https://ufal.mff.cuni.cz/iccl>

COLING 2025 is dedicated to providing a respectful, harassment-free environment for everyone. We do not tolerate harassment or bullying of any community member in any form. This does not only extend to attendees, but to anyone who chooses to become involved in the larger COLING 2025 community of organizers, volunteers and participants through events or interactions.

Harassment includes offensive verbal/electronic comments related to personal characteristics or choices, sexual images or comments in public or online spaces, deliberate intimidation, bullying, stalking, following, harassing photography or recording, sustained disruption of talks, IRC chats, electronic meetings, physical meetings or other events, inappropriate physical contact, or unwelcome sexual attention. Participants asked to stop any harassing or bullying behavior are expected to comply immediately.

If a participant engages in harassing behavior, representatives of COLING 2025 may take reasonable action they deem appropriate, including warning the offender, expulsion from mailing lists, discussion boards and other electronic communication channels or expulsion from the event to resolve the issue.

If you are being harassed, notice that someone else is being harassed, or have any other concerns, please act to intercede or ask for help from any member of the Local Organization Committee. Please be assured that your concerns will be kept in strict confidence, and we will consult with you on any actions taken.

This Code of Conduct has been adapted from the PyBCN one and is licensed under a Creative Commons Attribution-Share Alike 3.0 Unported license.

Meal Info

Breakfast (Jan. 19 - 24)

- Not Provided: However, many hotels offer complimentary breakfast options.

Breaks (Jan. 19 - 24)

- Coffee, Tea, and Light Bites: Available in both morning and mid-afternoon.

Lunch (Jan. 19 - 24)

- On Your Own: Lunch is not provided, so please explore local options.

Dinner (Jan. 19 - 24)

- On Your Own: Dinner is not provided, but there are many cafes and restaurants within walking distance.

Social Events

- Welcome Reception
 - Date: January 20, 19:00 - 21:00
 - Location: Aloft Hotel Splash Garden
 - Details: Light canapés and a drink ticket will be provided.
- Social Excursion Dinner in the Desert
 - Date: January 23, 19:00 - 22:30
 - Details: A BBQ buffet featuring local and international dishes, including grilled meats, vegetarian options, dessert, and refreshing drinks.

4

Welcome Reception

Welcome Reception – Join Us at Splash Garden, Aloft Abu Dhabi!



Date: **Monday Evening - Jan 20th**

Time:

18:00 - 20:30 – Registration Check-in

19:00 - 21:00 – Welcome Reception

Location: **Splash Garden, Aloft Abu Dhabi (Pool Level)**

Address: **Aloft Abu Dhabi, Al Khaleej Al Arabi Street, Abu Dhabi, UAE**

Kick off the event at the vibrant Splash Garden at Aloft Abu Dhabi! Enjoy light canapés and a drink ticket while mingling with fellow attendees in a relaxed, lively atmosphere. It's the perfect way to network and celebrate the start of our exciting week ahead. We can't wait to see you there!

Admission to the welcome reception is included in the main conference in-person registration. Additional tickets can be bought on site.

Additional Welcome Reception Ticket - \$75.

Welcome Reception - Under 10 - \$25.

Social Events – Offsite Excursion

COLING 2025 is thrilled to offer attendees a range of exciting social events and excursions in Abu Dhabi, providing the perfect mix of culture, history, and adventure. As part of the full conference registration, the Social Excursion Dinner in the Desert is included, and additional excursions are available for those interested in exploring the region in greater depth.

Social Excursion Dinner in the Desert

Date: **January 23, 19:00 - 22:30**

Details: Enjoy an unforgettable BBQ buffet under the stars in the desert, accompanied by refreshing beverages. This event is included with your full conference registration and promises to be a highlight of your COLING 2025 experience.

Excursion Options (Included for Full Conference Registration)

We are excited to offer 3 unique excursion options this year, each designed to cater to different interests:

1. For History Enthusiasts:
 - **Abrahamic House & Louvre Abu Dhabi:** Explore the rich cultural heritage of the region with a visit to the Abrahamic House, a symbol of unity, and the world-renowned Louvre Abu Dhabi, showcasing masterpieces from diverse cultures and historical periods.
2. For Architecture and Culture Lovers:
 - **Grand Mosque & Qasr Al Watan (Presidential Palace):** Discover the awe-inspiring beauty of the Sheikh Zayed Grand Mosque, followed by a visit to the magnificent Qasr Al Watan, the Presidential Palace, where you will experience the opulence and grandeur of Abu Dhabi's heritage.
3. For Adventure Seekers:
 - **Dune Bashing, Sand Surfing, and Camel Riding:** For those seeking adrenaline, embark on an exciting adventure through the desert with dune bashing and sand surfing, followed by a traditional camel ride. The excursion concludes with the unforgettable Social Excursion Dinner in the Desert.

Ticket Pricing for Additional Guests

- Social Event Ticket (Additional Guest): \$150
- Social Event Ticket (Under 10): \$75
- Social Event Ticket (Exhibitor): \$150

Tickets for these exciting excursions and social events can be purchased on-site. Be sure to book early, as spots are limited!

Social Excursion Option 1: Grand Mosque and Presidential Palace Afternoon Tour

1. Sheikh Zayed Grand Mosque

The day begins with a visit to the Sheikh Zayed Grand Mosque, one of the largest and most beautiful



mosques in the world. This architectural marvel is known for its stunning white marble exterior, grand prayer halls, and impressive minarets. Visitors will have the opportunity to explore the mosque's intricate interiors, including the largest hand-knotted carpet in the world, crystal chandeliers, and the peaceful courtyards. The mosque can accommodate over 40,000 worshippers and is a masterpiece of Islamic architecture, blending traditional and modern design elements.

Key highlights

- Majestic prayer halls
- Grand chandeliers and domes
- Expansive courtyards and reflective pools
- Impressive Islamic calligraphy and marble work

2. Qasr Al Watan (The Presidential Palace)

Next, the tour takes you to Qasr Al Watan, the presidential palace of the UAE. This magnificent palace



showcases the rich culture and heritage of the Emirates. Visitors will tour the public areas, including the opulent Great Hall, stunning gardens, and exhibitions detailing the UAE's history, governance, and future. The palace is an architectural and cultural wonder, with intricate designs and a breathtaking central dome that is among the largest in the world.

Key highlights

- Ornate designs and artwork reflecting UAE's history
- Magnificent Great Hall with its impressive dome
- Beautifully landscaped gardens
- Insight into the leadership and vision of the UAE

Social Excursion Option 2: Abrahamic House and Louvre Abu Dhabi Afternoon Tour

1. Abrahamic Family House

The tour begins with a visit to the Abrahamic Family House, an iconic and historic complex that celebrates



the shared values of the three Abrahamic religions: Islam, Christianity, and Judaism. Located on Saadiyat Island, this project is a symbol of peace, coexistence, and interfaith dialogue. The complex includes a mosque, a church, and a synagogue, each designed with stunning architecture that reflects the individual traditions while promoting unity. Visitors can explore the grounds, learn about the spiritual significance of each building, and appreciate the messages of tolerance and mutual respect.

Key highlights

- Explore the mosque, church, and synagogue, each with unique architectural styles.
- Learn about the interfaith dialogue and the cultural importance of religious unity.
- Reflect on the harmonious message of the Abrahamic religions.

2. Louvre Abu Dhabi

Next, you'll visit the Louvre Abu Dhabi, one of the most famous museums in the region. Located on



Saadiyat Island, this museum is a cultural beacon showcasing art, history, and civilizations from across the globe. The museum is renowned for its stunning architecture, including a massive dome that creates a mesmerizing light effect, often referred to as the “rain of light.” Inside, the Louvre displays a vast collection of artworks from ancient civilizations to contemporary pieces, including paintings, sculptures, and artifacts from various cultures and time periods. The museum promotes cultural exchange and the sharing of human heritage.

Key highlights

- A breathtaking collection of art and artifacts spanning different cultures and epochs.
- The iconic dome and the “rain of light” effect.

- Masterpieces from renowned artists such as Leonardo da Vinci, Van Gogh, and Monet.
- Rotating exhibits and installations that celebrate global heritage.

Social Excursion Options 3: Sand Dunes & 4x4 Desert Safari

1. Sand Dunes & 4x4 Desert Safari

The tour begins with an exciting 4x4 desert safari as guests are driven into the vast, golden dunes of the



Abu Dhabi desert. The ride is thrilling, with expert drivers navigating the towering sand dunes, providing a mix of dune bashing and scenic views. As you travel deeper into the desert, you'll have the chance to experience the striking beauty of the undulating sand dunes, which stretch as far as the eye can see.

Key highlights

- Exciting dune bashing in a 4x4 vehicle.
- Stunning views of the vast desert landscape.
- Thrill of riding over high dunes and through valleys.

2. Camel Riding

Next, you'll experience a more traditional desert mode of transport with a camel ride. Camels, known as



the “ships of the desert,” offer a unique and relaxing way to explore the landscape. As you gently sway atop the camel's back, you'll feel transported to an earlier time when camels were the main means of crossing the desert.

Key highlights

- Enjoy a camel ride through the dunes, offering a slower-paced experience.
- Capture stunning desert scenery from the back of the camel.
- Learn about the cultural significance of camels in Bedouin life.

3. Sand Surfing

For an additional adrenaline rush, the tour includes the opportunity to try sand surfing. You'll receive a sandboard and be guided to the perfect dune where you can attempt to glide down the sand slopes.



Whether you're an experienced surfer or a beginner, the thrill of gliding down the desert dunes is a unique and exhilarating experience.

Key highlights

- Sand surfing down the dunes for a fun adventure.
- Experience the sensation of surfing on soft, golden sand.
- Great for both beginners and those looking for a challenge.

Boxed Lunch As part of the tour, guests will enjoy a boxed lunch perfect for a quick and convenient meal between sightseeing stops.

Transport & Convenience The tour includes bus transfers, ensuring a smooth and comfortable journey between each destination, allowing guests to focus on enjoying the experience without worrying about logistics.

Dinner All Social Excursions meet up in the Desert for dinner under the Stars. As the evening sets in, the adventure transitions to the serene desert landscape. The experience includes a traditional desert camp, where a lavish dinner awaits under the starlit sky. The dinner spread includes a BBQ buffet of local and international dishes, from grilled meats to vegetarian options, along with dessert and refreshing drinks.

Entertainment As you enjoy your meal, the evening is further enhanced by live entertainment. Guests will be treated to a variety of traditional performances, including belly dancing, the colorful tanoura dance, and Arabic music, all contributing to the cultural atmosphere of the desert camp. The lively performances add to the enchantment of dining under the stars.



The 31st International
Conference on Computational
Linguistics

COLING Social Excursions

Reminder to bring your Name Badge for
access on the bus and at the desert

Busses start departing at 12:30 from ADNEC



Last Bus departs The Desert at 21:30

Keynotes

Learning a Language Like Infants do: Results and Challenges for Developmentally Inspired NLP

Emmanuel Dupoux

École des Hautes Études en Sciences Sociales (EHESS), Paris, France



Tuesday, Jan 21, 2025 - Room: Conference Hall A - Time: 09:30-10:30

Abstract: Instead of building AI systems that match human adults performance on tasks of interest, why not building an “AI child” that will be able to learn autonomously any task? This rather old idea has been notoriously difficult to implement, yet progress in machine learning and ecological datasets of parent-child interaction put us today in a good position to get a stab at it. We first describe the general conditions of autonomous language learning in the human child (continuous, scarce, noisy, multimodal, interactive data; fast, stable, overlapping learning curves), and propose methodological principles to compare child and machine learning abilities back to back. We then present some first results on (textless) speech language models showing a 3 to 5 orders of magnitude gap in sample efficiency (in favor of human children) and discuss competing hypotheses about what children have that current ai systems are lacking that could explain this large performance gap.

Bio: Emmanuel Dupoux is a Professor of Cognitive Psychology at the École des Hautes Études en Sciences Sociales (EHESS) in Paris. He earned his Ph.D. in Cognitive Psychology from EHESS in 1989, focusing on the mechanisms and representations that enable infants to acquire language and become cognitively functional within their culture. His research expertise lies in cognitive development, psycholin-

guistics, language acquisition, cognitive modeling, and machine learning. He specializes in studying early language acquisition, phonological ‘deafnesses’ in speech perception, and the development of social cognition. He also investigates how machine learning and artificial intelligence can provide quantitative models of processing and learning in infants. Throughout his career, he has received several awards and honors, including an Advanced ERC grant and the organization of the Zero Resource Speech Challenge (2015, 2017, 2019) and the Intuitive Physics Benchmark (2019).

Word Meaning, in Computational Linguistics and Beyond

Katrin Erk

University of Texas at Austin, USA



Wednesday, Jan 22, 2025 - Room: Conference Hall A - Time: 16:00-17:00

Abstract: What kinds of things are word meanings, and how can we model them computationally? In the time of distributional models, the study of polysemy structure always seemed just out of reach. Now, recent language models provide exhilarating new prospects for studying lexical meaning: Looking at embeddings, we can finally see usage groups, both similar to and intriguingly different from dictionary senses, especially in the cultural traces or “story traces” that they show. With prompting, we can get at more complex structures of meaning, like properties of frames and narratives. But there is so much to be figured out first. We have many promising techniques, but we don’t yet have reliable best practices and tools for analyzing language models for lexical meaning. It is also still difficult to distinguish the signal from the noise: When the picture that language models show us of word meaning diverges from how we humans organize dictionaries, then which parts are important facets of meaning that we have overlooked, and which parts are just peculiarities of the computational system? To make progress here, we also need to further develop our theories of the lexicon.

Bio: Katrin Erk is a Professor of Linguistics and Computer Science at the University of Texas at Austin. She earned her Ph.D. from Saarland University in Germany in 2002, focusing on tree description languages and ellipsis. Her research expertise lies in computational linguistics, particularly in semantics. She specializes in developing distributed, flexible approaches to describing word meaning and integrating them with representations at the sentence or discourse level. Her work includes studying flexible representations of word meaning constrained by context and exploring frameworks that draw inferences based on sentence structure and word meanings. She also investigates narrative schemas and their influence on word meaning and inference. In October 2024, it was announced that she will join the University of Massachusetts Amherst in September 2025, holding a joint position in the Department of Linguistics and the Manning College of Information and Computer Sciences. Throughout her career, she has received several awards and honors, including a CSLI Fellowship at Stanford in 2017 and a Google Faculty Research Award in 2018.

Towards Linguistically and Culturally Inclusive LLMs

Partha Talukdar
Google DeepMind, India



Friday, Jan 24, 2025 - Room: Conference Hall A - Time: 09:00-10:00

Abstract: Large Language Models (LLMs) have seen tremendous progress over the last few years with increasing adoption across the globe. Even though there are more than 7000 languages in the world, LLMs are currently usable in only a handful of them. Moreover, as LLMs get used across geographies, there is a need for them to become adaptive to regional cultural nuances and local norms. These raise interesting research challenges in language and culture, especially due to the limited availability of representative data and evaluations. In this talk, I shall present an overview of our research in this promising area of inclusive LLMs. I shall talk about Project Vaani where the goal is to capture and make available the speech landscape of India using a unique geo-anchored approach, modular approaches such as CALM to increase the scope of LLMs using composition, and benchmarks such as CUBE to evaluate the cultural knowledge of LLMs.

Bio: Partha is a Researcher at Google DeepMind India where he leads the Languages group. He is also a Faculty Member at IISc Bangalore. Previously, Partha was a Postdoctoral Fellow in the Machine Learning Department at Carnegie Mellon University. He received his PhD (2010) from the University of Pennsylvania. Partha is broadly interested in making AI more inclusive with a view towards benefiting a broader part of the world population. Partha is a recipient of several awards, including an Outstanding Paper Award at ACL 2019 and the ACM India Early Career Award 2022. He is also a Fellow of the Indian National Academy of Engineering. He is a co-author of a book on Graph-based Semi-Supervised Learning. Homepage: <https://parthatalukdar.github.io/>



Birds-of-a-Feather Group Discussions

January 21, 2025

Session 3: Arabic NLP

Organizer: **Nizar Habash**

Time: **14:00 - 15:30** - Location: **Hall B Room D**

Description: This Birds-of-a-Feather session is designed for researchers and developers working on or interested in Arabic NLP. Organized by a team affiliated with SIGARAB, the ACL Special Interest Group on Arabic Natural Language Processing, the session focuses on social networking and engaging discussions on timely and relevant topics.

Session 4: NLP for Social Good

Organizer: **Daryna Dementieva**

Time: **16:00 - 17:30** - Location: **Hall B Room D**

Description: While LLMs and other NLP/AI advancements present us with tremendous opportunities, they also demand greater responsibility from us as NLP researchers. One key responsibility is to bridge the gap between technology and real-world societal challenges, considering how AI/NLP can effectively address them. At our BoF session, we will discuss pressing issues—i.e. digital violence, fake news and propaganda detection (and feel free to propose your own topic). We will discuss both successful and less effective cases of NLP application to understand what drives these outcomes. Our aim is to create a space for researchers worldwide to connect and collectively envision a brighter future, using NLP for the greater good.

January 22, 2025

Session 6: Language Technology for Crisis Preparedness and Response

Organizer: **Belu Ticona**

Time: **11:00 - 12:30** - Location: **Hall B Room D**

Description: How can we use language technology to prepare for the next crisis? In a world with more natural disasters and socio-political conflicts, communication is crucial so that victims can share their needs (in their languages/dialects), organizations can provide humanitarian aid, and broadcast urgent in-

formation. In this BoF, we will discuss challenges and possible solutions through a role-playing activity. We invite anyone interested in low-resource languages, machine translation, multicultural NLP, climate change, or crisis and language in general.

Session 7: Southeast Asian NLP

Organizer: **Jan Christian Blaise Cruz**

Time: **14:00 - 15:30** - Location: **Hall B Room D**

Description: Small group discussions and community-driven talks around the current challenges, directions, and opportunities in Southeast Asian NLP. The BoF is organized by co-founders of the ACL SIGSEA and SEACrowd. We aim to host one BoF session in every major NLP conference after the success of the ones hosted in ACL 2024 and EMNLP 2024.

January 23, 2025

Session 9: Supporting Mental Health Through NLP-Powered Conversational Agents

Organizer: **Sadegh Jafari**

Time: **09:00 - 10:30** - Location: **Hall B Room D**

Description: We will explore how advanced language models and classification techniques can detect stress disorders, analyze emotions, and provide personalized assistance to users. Attendees will have the opportunity to discuss challenges, ethical considerations, and future directions for using NLP in mental health, fostering collaboration between researchers, practitioners, and technologists passionate about making a positive impact on mental well-being.

Session 10: Multilingual and Multimodal Cultural Inclusivity in LLMs

Organizer: **Firoj Alam**

Time: **11:00 - 12:30** - Location: **Hall B Room D**

Description: The Birds of a Feather session on "Multilingual and Multimodal Cultural Inclusivity in LLMs" aims to foster a collaborative dialogue on the challenges and opportunities of creating LLMs that are inclusive of diverse languages, cultures, and modalities. As LLMs increasingly influence global communication and understanding, ensuring they represent and respect cultural nuances, underrepresented languages, and multimodal interactions is critical. This session will explore current advancements, challenges in addressing linguistic and cultural biases, and strategies for incorporating multilingual and multimodal datasets into LLM training. Participants will discuss the ethical, technical, and societal implications of these efforts, sharing ideas and solutions to make AI more inclusive and equitable across diverse linguistic and cultural landscapes.

January 24, 2025

Session 12: Exploring Public Sentiment on AI

Organizer: **Rushil Thareja**

Time: **10:30 - 12:00** - Location: **Hall B Room D**

Description: With the widespread adoption of AI across various aspects of life, productivity has increased, and many find life simpler and more convenient. Yet, a growing segment of society views AI with skepticism or negativity, often perceiving it as unnecessary, harmful, or driven by capitalist motives. This sentiment raises a critical question: if AI is meant to improve lives, why do so many feel otherwise? The goal of this session is for AI researchers to come together to understand these perceptions, challenge assumptions about their root causes, and rethink how AI is designed and implemented. By aligning AI with people's needs and values, we can ensure it becomes a truly positive force that works hand in hand with society.

Conference Overview

Saturday, January 18, 2025

Breakfast and dinner are not provided.
The Arabic NLP Winter School will offer Lunch

| | | |
|---------------|---|------------------|
| 08:50 - 21:00 | Arabic NLP Winter School - Day 1 <i>Detailed schedule on school's website</i> | MBZUAI (Offsite) |
| 08:00 - 09:00 | Arabic NLP Winter School Registration | MBZUAI (Offsite) |
| 14:00 - 19:00 | COLING Registration | ADNEC |

Sunday, January 19, 2025

Breakfast, lunch and dinner are not provided.
See detailed schedules under Workshop/Tutorials/Arabic NLP Winter School

| | |
|---------------|--|
| 07:30 - 16:30 | Registration |
| 10:30 - 11:00 | Break (Coffee, Tea & Light Bites) |
| 15:30 - 16:00 | Break (Coffee, Tea & Light Bites) |

Monday, January 20, 2025

Breakfast, lunch and dinner are not provided.
See detailed schedules under Workshop/Tutorials

| | | |
|---------------|---|-------|
| 07:30 - 17:30 | Registration <i>For all Registered Full, Main, or Welcome Reception addon's</i> | ADNEC |
|---------------|---|-------|

10:30 - 11:00 **Break** (Coffee, Tea & Light Bites)

15:30 - 16:00 **Break** (Coffee, Tea & Light Bites)

18:00 - 20:30 **Welcome Reception Registration**

Aloft Splash Garden

19:00 - 21:00 **Welcome Reception**

Aloft Splash Garden

Welcome drink ticket, light bites. Registered attendees only

Tuesday, January 21, 2025

Breakfast, lunch and dinner are not provided.

07:30 - 16:30 **Registration**

09:00 - 09:30 **Session 1: Plenary – Opening Session**

09:30 - 10:30 **Invited Speaker: Emmanuel Dupoux**

"Learning a Language Like Infants Do: Results and Challenges for Developmentally Inspired NLP"

10:30 - 11:00 **Break** (Coffee, Tea & Light Bites)

11:00 - 12:30 **Session 2: Group Discussions (A)**

ORALS

Multilinguality and Machine Translation 1

Hall B (A)

Semantics and Reasoning 1

Hall B (B)

Natural Language Generation and Summarization 1

Hall B (C)

NLP Applications 1

Capital Suite 7

POSTER/DEMOS

Atrium

NLP Applications

Information Extraction and Retrieval

Cognitive Modeling and Language Acquisition

Dialogue and Conversational Interaction

BoF

TBA

Hall B (D)

12:30 - 14:00 **Lunch Break**

All attendees on own

14:00 - 15:30 **Session 3: Group Discussions (B)**

ORALS

Discourse, Phonology and Syntax

Hall B (A)

Machine Learning

Hall B (B)

Interpretability and Explainability

Hall B (C)

Language Resources and Evaluation 1

Capital Suite 7

POSTER/DEMOS

Atrium

Semantics and Reasoning

NLP Applications

Multilinguality, Machine Translation and Low-Resource Languages
 Natural Language Generation and Summarization

BoF

Arabic NLP Hall B (D)

15:30 - 16:00 **Break (Coffee, Tea & Light Bites)**

16:00 - 17:30 **Session 4: Group Discussions (C)**

ORALS

Cognitive Modelling Hall B (A)
 Information Extraction and Retrieval 1 Hall B (B)
 Dialogue and Conversational Interaction 1 Hall B (C)
 NLP Applications 2 Capital Suite 7

POSTER/DEMOS

Discourse, Phonology and Syntax Atrium
 Interpretability and Explainability
 Language Resources and Evaluation
 Machine Learning

BoF

NLP for Social Good Hall B (D)

Wednesday, January 22, 2025

Breakfast, lunch and dinner are not provided.

08:30 - 16:30 **Registration**

09:00 - 10:30 **Session 5: Group Discussions (D)**

ORALS

Multimodal NLP 1 Hall B (A)
 Information Extraction and Retrieval 2 Hall B (B)
 Industry 1 Hall B (C)
 Low-Resource Languages 1 Capital Suite 7

POSTER/DEMOS

NLP Applications Atrium
 Ethics, Bias and Fairness
 Multilinguality, Machine Translation and Low-Resource Languages

BoF

TBA Hall B (D)

10:30 - 11:00 **Break (Coffee, Tea & Light Bites)**

11:00 - 12:30 **Session 6: Group Discussions (E)**

ORALS

Ethics, Bias and Fairness 1 Hall B (A)

Multilinguality and Machine Translation 2 Hall B (B)
Industry 2 Hall B (C)
NLP Applications 3 Capital Suite 7

POSTER/DEMOS

Language Modeling
Semantics and Reasoning
Language Resources and Evaluation
Sentiment Analysis

Atrium

BoF

Language Technology for Crisis Preparedness and Response Hall B (D)

12:30 - 14:00 **Lunch Break** All attendees on own

14:00 - 15:30 **Session 7: Group Discussions (F)**

ORALS

Semantics and Reasoning 2
Language Modeling 1
Industry 3
Language Resources and Evaluation 2

Hall B (A)
Hall B (B)
Hall B (C)
Capital Suite 7

POSTER/DEMOS

Information Extraction and Retrieval
Multimodal NLP
Machine Learning

Atrium

BoF

Southeast Asian NLP

Hall B (D)

15:30 - 16:00 **Break (Coffee, Tea & Light Bites)**

16:00 - 17:00 **Session 8: Plenary**

Invited Speaker: Professor Katrin Erk

"Word Meaning, in Computational Linguistics and Beyond"

Thursday, January 23, 2025

Breakfast, lunch and dinner are not provided.

08:30 - 16:30 **Registration**

09:00 - 10:30 **Session 9: Group Discussions (G)**

ORALS

Low-Resource Languages 2
Semantics and Reasoning 3
Language Resources and Evaluation 3
NLP Applications 4

Hall B (A)
Hall B (B)
Hall B (C)
Capital Suite 7

POSTER/DEMOS

Atrium

Industry

BoF

Supporting Mental Health Through NLP-Powered Conversational Agents Hall B (D)

10:30 - 11:00 **Break (Coffee, Tea & Light Bites)**

11:00 - 12:30 **Session 10: Group Discussions (H)**

ORALS

| | |
|-------------------------------------|-----------------|
| Multimodal NLP 2 | Hall B (A) |
| Sentiment Analysis 1 | Hall B (B) |
| Ethics, bias and fairness 2 | Hall B (C) |
| Language Resources and Evaluation 4 | Capital Suite 7 |

POSTER/DEMOS

| | |
|---|--------|
| Multilinguality, Machine Translation and Low-Resource Languages | Atrium |
| Language Resources and Evaluation | |
| Semantics and Reasoning | |
| NLP Applications | |

BoF

Multilingual and Multimodal Cultural Inclusivity in LLMs Hall B (D)

12:30 - 22:30 **Social Excursions**

1 – A half-day tour of the Grand Mosque and the Palace followed by a dinner in the desert under the stars, offers a truly unique and memorable experience.
 2 – A half-day tour of the Abrahamic Family House and the Louvre Abu Dhabi, followed by a desert dinner under the stars, offers a perfect blend of cultural exploration and adventure.
 3 – A half-day desert tour in Abu Dhabi that includes sand dunes, camel riding, sand surfing, a boxed lunch, and dinner under the stars offers an exhilarating and immersive experience into the heart of the Arabian desert.

Friday, January 24, 2025

Breakfast, lunch and dinner are not provided.

08:30 - 16:30 **Registration**

09:00 - 10:00 **Session 11: Plenary**

Invited Speaker: Professor Partha Talukdar
"Towards Linguistically and Culturally Inclusive LLMs"

10:00 - 10:30 **Break (Coffee, Tea & Light Bites)**

| | | |
|---------------|---|-----------------|
| 10:30 - 12:00 | Session 12: Group Discussions (I) | |
| | ORALS | |
| | Sentiment Analysis 2 | Hall B (A) |
| | Natural Language Generation and Summarization 2 | Hall B (B) |
| | Dialogue and Conversational Interaction 2 | Hall B (C) |
| | Multilinguality and Machine Translation 3 | Capital Suite 7 |
| | POSTER/DEMOS | Atrium |
| | NLP Applications | |
| | Language Resources and Evaluation | |
| | Language Modeling | |
| | Information Extraction and Retrieval | |
| | BoF | |
| | Exploring Public Sentiment on AI | Hall B (D) |
| 12:00 - 13:30 | Lunch Break | |
| 13:30 - 15:00 | Session 13: Group Discussions (J) | |
| | ORALS | |
| | Language Resources and Evaluation 5 | Hall B (A) |
| | Language Modeling 2 | Hall B (B) |
| | Information Extraction and Retrieval 3 | Hall B (C) |
| | Multimodal NLP 3 | Capital Suite 7 |
| | POSTER/DEMOS | Atrium |
| | Natural Language Generation and Summarization | |
| | Dialogue and Conversational Interaction | |
| | Sentiment Analysis | |
| | Multimodal NLP | |
| | BoF | |
| | TBA | Hall B (D) |
| 15:00 - 15:30 | Break (Coffee, Tea & Light Bites) | |
| 15:30 - 16:30 | Session 14: Awards | |
| 16:30 - 17:00 | Closing Remarks | |

Tutorials Overview

Tutorials Overview: Sunday, January 19, 2025

T1: Speculative Decoding for Efficient LLM Inference

Organizers: Heming Xia, Cunxiao Du, Yongqi Li, Qian Liu, and Wenjie Li

Time: 09:00-12:30

Location: Capital Suite 7

T2: From Theory to Practice: A Hands-on Tutorial in Explainable NLP

Organizers: Wafa Abdullah Alrajhi, Nourah Alangari, and Hend Al-Khalifa

Time: 09:00-12:30

Location: Online Only

T3: Hands-On Tutorial: Labeling with LLM and Human-in-the-Loop

Organizers: Ekaterina Artemova, Akim Tsvigun, Dominik Schlechtweg, Natalia Fedorova, Sergei Tilga, Boris Obmoroshev, and Konstantin Chernyshev

Time: 14:00-17:30

Location: Capital Suite 5

T4: LLMs in Education: Novel Perspectives, Challenges, and Opportunities

Organizers: Bashar Alhafni, Sowmya Vajjala, Stefano Banno, Kaushal Kumar Maurya, and Ekaterina Kochmar

Time: 14:00-17:30

Location: Capital Suite 7

T5: EduRAG: Crafting Clever Educational Chatbots and Advanced QA Systems with Retrieval-Augmented Generation

Organizers: Noorhan Abbas and Saad Ezzini

Time: 14:00-17:30

Location: Online Only

Tutorial Overviews: Monday, January 20, 2025

T6: Connecting Ideas in ‘Lower-Resource Scenarios: NLP for National Varieties, Creoles and Other Low-resource Scenarios

Organizers: Aditya Joshi, Diptesh Kanojia, Heather Lent, Hour Kaing, and Haiyue Song

Time: 09:00-17:30

Location: Conference Hall B (C)

T7: Safety Issues for Generative AI

Organizers: Haonan Li, Xudong Han, Emad A. Alghamdi, Shom Lin, Monojit Choudhury, Jingfeng Zhang, Paul Rottger and Timothy Baldwin

Time: 09:00-12:30

Location: Capital Suite 7

T8: Hallucinative Foundation Models: Characterization, Quantification, Avoidance, and Mitigation

Organizers: Vipula Rawte, Aman Chadha, Amit Sheth and Amitava Das

Time: 09:00-12:30

Location: Online Only

T9: Bridging Linguistic Theory and AI: Usage-Based Learning in Humans and Machines

Organizers: Claire Bonial, Harish Tayyar Madabushi, Nikhil Krishnaswamy, and James Pustejovsky

Time: 14:00-17:30

Location: Capital Suite 7

Tutorials Details

Sunday, January 19, 2025

Tutorial 1: Speculative Decoding for Efficient LLM Inference

Organizers: Heming Xia, Cunxiao Du, Yongqi Li, Qian Liu and Wenjie Li

This tutorial presents a comprehensive introduction to Speculative Decoding (SD), an advanced technique for LLM inference acceleration that has garnered significant research interest in recent years. SD is introduced as an innovative decoding paradigm to mitigate the high inference latency stemming from autoregressive decoding in LLMs. At each decoding step, SD efficiently drafts several future tokens and then verifies them in parallel. This approach, unlike traditional autoregressive decoding, facilitates the simultaneous decoding of multiple tokens per step, thereby achieving promising 2x-4x speedups in LLM inference while maintaining original distributions. This tutorial delves into the latest techniques in SD, including draft model architectures and verification strategies. Additionally, it explores the acceleration potential and future research directions in this promising field. We aim for this tutorial to elucidate the current research landscape and offer insights for researchers interested in Speculative Decoding, ultimately contributing to more efficient LLM inference.

Date: Jan. 19, 2025

Time: 09:00-12:30

Location: Capital Suite 7

Tutorial 2: From Theory to Practice: A Hands-on Tutorial in Explainable NLP

Organizers: Wafa Abdullah Alrajhi, Nourah Alangari, and Hend Al-Khalifa

This tutorial aims to explore the intersection between XAI and the field of Natural Language Processing (NLP) by highlighting its significance for various tasks. By combining both theoretical and practical aspects of XAI, we provide a comprehensive overview of its definitions and review the state-of-the-art methodologies employed to explain NLP models. While previous tutorials have focused on establishing the theoretical foundations of the XAI field and providing essential concepts to the audience, this tutorial seeks to bridge the gap between theory and practice. We will discuss and demonstrate the most commonly used XAI techniques, the prominent libraries applicable to different models, and the limitations of each method. Furthermore, we will highlight the methods that can be utilized to visualize these models while performing specific NLP tasks.

Date: Jan. 19, 2025
Time: 09:00-12:30
Location: Online Only

Tutorial 3: Hands-On Tutorial: Labeling with LLM and Human-in-the-Loop

Organizers: Ekaterina Artemova, Akim Tsvigun, Dominik Schlechtweg, Natalia Fedorova, Sergei Tilga, Boris Obmoroshev and Konstantin Chernyshev

Training and deploying machine learning models relies on a large amount of human-annotated data. As human labeling becomes increasingly expensive and time-consuming, recent research has developed multiple strategies to speed up annotation and reduce costs and human workload: generating synthetic training data, active learning, and hybrid labeling. This tutorial is oriented toward practical applications: we will present the basics of each strategy, highlight their benefits and limitations, and discuss in detail real-life case studies. Additionally, we will walk through best practices for managing human annotators and controlling the quality of the final dataset. The tutorial includes a hands-on workshop, where attendees will be guided in implementing a hybrid annotation setup. This tutorial is designed for NLP practitioners from both research and industry backgrounds who are involved in or interested in optimizing data labeling projects.

Date: Jan. 19, 2025
Time: 14:00-17:30
Location: Capital Suite 5

Tutorial 4: LLMs in Education: Novel Perspectives, Challenges, and Opportunities

Organizers: Bashar Alhafni, Sowmya Vajjala, Stefano Banno, Kaushal Kumar Maurya and Ekaterina Kochmar

The role of large language models (LLMs) in education is an increasing area of interest today, considering the new opportunities they offer for teaching, learning, and assessment. This cutting-edge tutorial provides an overview of the educational applications of NLP and the impact that the recent advances in LLMs have had on this field. We will discuss the key challenges and opportunities presented by LLMs, grounding them in the context of four major educational applications: reading, writing, and speaking skills, and intelligent tutoring systems (ITS). This tutorial is designed for researchers and practitioners interested in the educational applications of NLP and the role LLMs have to play in this area. It is the first of its kind to address this timely topic.

Date: Jan. 19, 2025
Time: 14:00-17:30
Location: Capital Suite 7

Tutorial 5: EduRAG: Crafting Clever Educational Chatbots and Advanced QA Systems with Retrieval-Augmented Generation

Organizers: Noorhan Abbas and Saad Ezzini

Date: Jan. 19, 2025
Time: 14:00-17:30
Location: Online Only

Monday, January 20, 2025

Tutorial 6: Connecting Ideas in Lower-Resource Scenarios: NLP for National Varieties, Creoles, and Other Low-Resource Scenarios

Organizers: Aditya Joshi, Diptesh Kanojia, Heather Lent, Hour Kaing and Haiyue Song

Despite excellent results on benchmarks over a small subset of languages, large language models struggle to process text from languages situated in ‘lower-resource’ scenarios such as dialects/sociolects (national or social varieties of a language), Creoles (languages arising from linguistic contact between multiple languages) and other low-resource languages. This introductory tutorial will identify common challenges, approaches, and themes in natural language processing (NLP) research for confronting and overcoming the obstacles inherent to data-poor contexts. By connecting past ideas to the present field, this tutorial aims to ignite collaboration and cross-pollination between researchers working in these scenarios. Our notion of ‘lower-resource’ broadly denotes the outstanding lack of data required for model training - and may be applied to scenarios apart from the three covered in the tutorial.

Date: Jan. 20, 2025

Time: 09:00 - 17:30

Location: Conference Hall B (C)

Tutorial 7: Safety Issues for Generative AI

Organizers: Haonan Li, Xudong Han, Emad A. Alghamdi, Shom Lin, Monojit Choudhury, Jingfeng Zhang, Paul Rottger and Timothy Baldwin

This tutorial will provide an in-depth exploration of safety issues for generative AI models, covering a broad range of sub-topics, including a risk taxonomy, adversarial attack types, safety evaluation, defense mechanisms, red teaming for multi-modal models, and agentic AI. The goal is to brief attendees on the latest advancements and emerging trends in the field, enabling them to identify and mitigate vulnerabilities in generative AI systems. Participants will gain insights into recent research developments, open research questions, and novel research directions. This cutting-edge tutorial is designed for AI researchers, developers, and security professionals with a basic knowledge of red teaming and AI safety.

Date: Jan. 20, 2025

Time: 09:00-12:30

Location: Capital Suite 7

Tutorial 8: Hallucinative Foundation Models: Characterization, Quantification, Avoidance, and Mitigation

Organizers: Vipula Rawte, Aman Chadha, Amit Sheth and Amitava Das

In the fast-paced domain of Large Language Models (LLMs), the issue of hallucination is a prominent challenge. Despite continuous endeavors to address this concern, it remains a highly active area of research within the LLM landscape. Grasping the intricacies of this problem can be daunting, especially for those new to the field. This tutorial aims to bridge this knowledge gap by introducing the emerging realm of hallucination in LLMs. It will comprehensively explore the key aspects of hallucination, including benchmarking, detection, and mitigation techniques. Furthermore, we will delve into the specific constraints and shortcomings of current approaches, providing valuable insights to guide future research efforts for participants.

Date: Jan. 20, 2025

Time: 09:00-12:30

Location: Online Only

Tutorial 9: Bridging Linguistic Theory and AI: Usage-Based Learning in Humans and Machines

Organizers: Claire Bonial, Harish Tayyar Madabushi, Nikhil Krishnaswamy and James Pustejovsky

Usage-based theories of human language, such as Construction Grammar, have been compelling theoretical lenses through which to view and evaluate what LLMs know and understand of language because of the parallels between usage-based learning and the data-driven “learning” of pre-trained models. However, a key difference between a usage-based learning account for humans and that of LLMs is in embodiment and multimodality—for the most part, LLMs use text alone, whereas usage-based theories posit that each token of linguistic experience is stored with a wealth of experiential information that enriches the symbol through cross-modal association. Therefore, the first goal of this tutorial is to provide a summary of language acquisition and second language learning from a usage-based theoretical linguistic perspective. With this understanding of human usage-based learning, we will turn to evidence demonstrating the ways in which machine learning, primarily via large, pre-trained vision and language models, does and does not parallel human learning. The overarching goal of this is not to say that the two processes are similar or dissimilar in order to conclude that dissimilarity denotes inferiority (if the knowledge arrived at is the same, then it may not matter how it was learned). Rather, we explore the resulting differences in what is known and understood about the world, and take this as a starting point for considering how to supplement and improve natural language understanding (NLU), particularly for physically situated applications. Our target audience is those interested in the intersection of linguistic theory and NLU implementations, such as human-robot interaction.

Date: Jan. 20, 2025

Time: 14:00-17:30

Location: Capital Suite 7

11

Workshops Overview

Workshops Overview: Sunday, January 19, 2025

W1: **The 1st Workshop on NLP for Languages Using Arabic Script (AbjadNLP 2025)**
Time: 09:00-16:45
Location: Conf. Hall B Room D

W2: **CHIPSAL: Challenges in Processing South Asian Languages**
Time: 08:30-15:45
Location: Online Only

W3: **Context and Meaning: Navigating Disagreements in NLP Annotation (CM-ND-NLP)**
Time: 09:30-16:10
Location: Capital Suite 6 (Level 2)

W4: **New Horizons in Computational Linguistics for Religious Texts (Coling-Rel)**
Time: 09:00-19:00
Location: Capital Suite 10 (Level 2)

W5: **1st Workshop on Computational Humor (CompHum)**
Time: 08:45-17:40
Location: Online Only

W6: **COLING 2025 Workshop on Detecting AI Generated Content (GenAIDetect)**
Time: 09:50-16:30
Location: Capital Suite 9 (Level 2)

W7: **Workshop on Generative AI and Knowledge Graphs (GenAIK)**

Workshops Overview

Time: 08:45-17:40
Location: Capital Suite 8 (Level 2)

W8: The First Workshop on Multilingual Counterspeech Generation (MCG)
Time: Check online for Schedule
Location: Capital Suite 13 (Level 2)

W9: VarDial 2025: The Twelfth Workshop on NLP for Similar Languages, Varieties and Dialects
Time: 09:00-17:30
Location: Conf. Hall B Room A

Workshops Overview: Sunday-Monday, January 19-20, 2025

W10: The Joint Workshop of the 9th Financial Technology and Natural Language Processing (FinNLP), the 6th Financial Narrative Processing (FNP), and the 1st Workshop on Large Language Models for Finance and Legal
Time: 09:00-17:10
Location: Conf. Hall B Room B

Workshops Overview: Sunday-Monday, January 20, 2025

W11: The Second COLING Workshop on Bridging Neurons and Symbols for NLP and Knowledge Graph Reasoning (BNS-KGR)
Time: 08:50-18:00
Location: Conf. Hall B Room A

W12: 18th Workshop on Building and Using Comparable Corpora (BUCC)
Time: 09:15-17:45
Location: Capital Suite 9 (Level 2)

W13: The Fifth Celtic Language Technology Workshop (CLT)
Time: 09:00-13:25
Location: Online Only

W14: Workshop of Evaluation of Multimodal Generation (EvalMG)
Time: Check online for Schedule
Location: Capital Suite 10 (Level 2)

W15: IndoNLP: The First Workshop on Natural Language Processing for Indo-Aryan and Dravidian Languages
Time: 08:45-17:00
Location: Conf. Hall B Room D

W16: **The First Workshop on Language Models for Low-Resource Languages (LoResLM 2025)**

Time: 08:45-18:00

Location: Capital Suite 5 (Level 2)

W17: **The 1st International Workshop on Nakba Narratives as Language Resources (Nakba-NLP 2025)**

Time: 09:00-17:15 GMT

Location: Online Only

W18: **RegNLP-2025: The First Workshop on Regulatory NLP**

Time: 09:00-17:30

Location: Capital Suite 13 (Level 2)

W19: **The Second Workshop on South East Asian Language Processing (SEALP)**

Time: Check online for Schedule

Location: Online Only

W20: **SUMEval-2: The 2nd Workshop on Scaling Up Multilingual & Multi-Cultural Evaluation**

Time: 09:00-17:30

Location: Capital Suite 6 (Level 2)

W21: **The 4th Workshop on Arabic Corpus Linguistics (WACL-4)**

Time: 09:00-16:30

Location: Online Only

W22: **Writing Aids at the Crossroads of AI, Cognitive Science and NLP (WR-AI-CogS)**

Time: 09:00-18:30

Location: Capital Suite 8 (Level 2)

Workshop Details

Sunday, January 19, 2025

Workshop 1: The 1st Workshop on NLP for Languages Using Arabic Script (AbjadNLP 2025)

Organizers: Mo El-Haj, Hugh Paterson III, Saad Ezzini, Ignatius Ezeani, Mahum Hayat Khan, Muhammad Sharjeel, Sina Ahmadi, Cynthia Amol, Amal Haddad Haddad, Jaleh Delfani, Ruslan Mitkov, and Paul Rayson

AbjadNLP is dedicated to advancing innovation and gaining deeper insights into Natural Language Processing (NLP) for languages that use the Arabic script. Our primary focus is on Abjad and Ajami languages that utilise the Arabic script or its variations. Traditionally associated with Semitic languages, Abjad scripts represent consonants in every syllable. In contrast, Ajami scripts denote the alphabetic use of the Arabic script in various African contexts, representing non-Arabic languages. We are interested in research on languages that fall under the Abjad or Ajami categories that use the Arabic script or any variations of it.

<https://wp.lancs.ac.uk/abjad/>

Time: 09:00-16:45

Location: Conf. Hall B Room D

Workshop 2: CHiPSAL: Challenges in Processing South Asian Languages

Organizers: Kengatharaiyer Sarveswaran, Ashwini Vaidya, Bal Krishna Bal, Sana Shams, and Surendra-bikram Thapa

South Asia, consisting of Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka, is one of the most populous regions in the world, with approximately 1.97 billion people. Home to over 700 languages and around 25 major scripts, the region reflects a rich cultural and linguistic heritage. Additionally, more than 50 million South Asians live abroad. Despite this diversity, South Asian languages are underrepresented in language technology. Recent large language models (LLMs) include minimal data from this region, and the challenges in processing South Asian languages begin with encoding issues. While most scripts are encoded in Unicode, some applications may not render them correctly due to orthographic complexities, and language input remains a problem in the region. The linguistic intricacies of these languages, with multiple writing systems and long literary traditions, further complicate natural language processing tasks. Dialectal and cultural variations, as well as close language contact, add an extra level of complexity. Therefore, this workshop focuses on the challenges in processing South

Workshops Details

Asian languages, covering issues related to linguistic and cultural aspects, encoding and orthography, and resource constraints. By addressing these challenges, we aim to facilitate South Asian language processing with a focus on linguistic and cultural heritage.

<https://sites.google.com/view/chipsal/>

Time: 08:30-15:45

Location: Online Virtual Only

Workshop 3: Context and Meaning: Navigating Disagreements in NLP Annotation (CM-ND-NLP)

Organizers: Michael Roth and Dominik Schlechtweg

Disagreements among annotators pose a significant challenge in Natural Language Processing, impacting the quality and reliability of datasets and consequently the performance of NLP models. This workshop aims to explore the complexities of annotation disagreements, their causes, and strategies towards their effective resolution, with a focus on meaning in context.

<https://comedinlp.github.io/>

Time: 09:30-16:10

Location: Capital Suite 6 (Level 2)

Workshop 4: New Horizons in Computational Linguistics for Religious Texts (Coling-Rel)

Organizers: Majdi Sawalha, Sane Yagi, Faisal Alshargi, Abdallah Al-Shadifat, Ashraf Elnagar, Bayan Abu Shawar, and Noorhan Abbas

The COLING 2025 Workshop on New Horizons in Computational Linguistics for Religious Texts will be held with the 31st edition of COLING in 2025 in Abu Dhabi (UAE) (COLING 2025). This workshop invites researchers exploring the intersection of language technology and religious texts. This workshop aims to foster discussion on cutting-edge applications of Natural Language Processing (NLP) to religious texts, including: (i) Analyzing faith-defining canons and authoritative interpretations, (ii) Extracting insights from sermons, liturgy, prayers, and poetry, and (iii) Leveraging Large Language Models for novel research avenues. We will explore the potential of NLP to unlock new understandings of religious traditions and chart the future of this exciting research area. The workshop welcomes researchers from computational linguistics, digital humanities, and related fields.

<https://668b0f080fa32.site123.me/>

Time: 09:00-19:00

Location: Capital Suite 10 (Level 2)

Workshop 5: 1st Workshop on Computational Humor (CompHum)

Organizers: Christian F. Hempelmann, Julia Rayz, Tiansi Dong, and Tristan Miller

The impressive recent progress in generative AI has enabled new approaches to complex tasks that can be thought of as essentially human, such as producing or understanding humor. For its part, humor research has become a mature, interdisciplinary field, both in theoretical advances across disciplines such as psychology, linguistics, and sociology, and in a breadth of purview and empirical support. Our workshop aims to foster further work on modeling the processes of humor with current methods in computational linguistics and natural language processing, against the theoretical backdrop of humor research and with reference to relevant corpora of textual, visual, and multimodal materials.

<https://chum2025.github.io/>

Time: 08:45-17:40

Location: Online Virtual Only

Workshop 6: COLING 2025 Workshop on Detecting AI Generated Content (GenAIDetect)

Organizers: Firoj Alam, Preslav Nakov, Nizar Habash, Iryna Gurevych, Shammur Chowdhury, Artem Shelmanov, Yuxia Wang, Ekaterina Artemova, Mucahid Kutlu, and George Mikros

This workshop aims to address the critical aspects of detecting generative content, encompassing text, images, and multimodal interactions. The emergence of advanced generative models has made it increasingly difficult to distinguish between human-written and machine-generated content, raising concerns in areas such as misinformation, copyright infringement, and the authenticity of digital media.

<https://genai-content-detection.gitlab.io/>

Time: 09:50-16:30

Location: Capital Suite 9 (Level 2)

Workshop 7: Workshop on Generative AI and Knowledge Graphs (GenAIK)

Organizers: Genet Asefa Gesese, FIZ Karlsruhe, Harald Sack, Heiko Paulheim, Albert Meroño-Peñuela, and Lihu Chen

In recent times, Knowledge Graphs (KGs) have attracted considerable attention for their ability to represent structured and interconnected information, and adopted by many companies in various domains. KGs represent knowledge by depicting relationships between entities, known as facts, usually based on formal ontological models. Consequently, they enable accuracy, decisiveness, interpretability, domain-specific knowledge, and evolving knowledge in various AI applications. The intersection between GenAI and KG has ignited significant interest and innovation in Natural Language Processing (NLP). For instance, by integrating LLMs with KGs during pre-training and inference, external knowledge can be incorporated for enhancing the model's capabilities and improving interpretability. When integrated, they offer a robust approach to problemsolving in diverse areas such as information enrichment, representation learning, conversational AI, cross-domain AI transfer, bias, content generation, and semantic understanding. This workshop, aims to reinforce the relationships between Deep Learning, Knowledge Graphs, and NLP communities and foster inter-disciplinary research in the area of GenAI.

<https://genetasefa.github.io/GenAIK2025/>

Time: 08:45-17:40

Location: Capital Suite 8 (Level 2)

Workshop 8: The First Workshop on Multilingual Counterspeech Generation (MCG)

Organizers: Rodrigo Agerri, Helena Bonaldi, Marco Guerini, María Teresa Martín-Valdivia, Arturo Montejo-Ráez, Aitor Soroa, María Estrella Vallecillo-Rodríguez, and Irune Zubiaga

While interest in automatic approaches to Counterspeech generation has been steadily growing, including studies on data curation, detection, and generation, the large majority of the published experimental work on automatic Counterspeech generation has been carried out for English. This is due to the scarcity of both non-English manually curated training data and to the crushing predominance of English in the generative Large Language Models (LLMs) ecosystem. A workshop on exploring Multilingual Counterspeech Generation is proposed to promote and encourage research on multilingual approaches for this challenging topic.

Workshops Details

<https://sites.google.com/view/multilang-counterspeech-gen/>

Time: Check online for Schedule

Location: Capital Suite 13 (Level 2)

Workshop 9: VarDial 2025: The Twelfth Workshop on NLP for Similar Languages, Varieties and Dialects

Organizers: Yves Scherrer, Tommi Jauhiainen, Nikola Ljubešić, Preslav Nakov, Jörg Tiedemann, and Marcos Zampieri

VarDial is a well-established series of workshops promoting a forum for scholars working on a range of topics related to the study of diatopic language variation from a computational perspective. The workshop deals with computational methods and language resources for closely related languages, language varieties, and dialects. We welcome papers dealing with one or more of the following topics: (i) Corpora, resources, and tools for similar languages, varieties and dialects; (ii) Adaptation of tools (taggers, parsers) for similar languages, varieties and dialects; (iii) Evaluation of language resources and tools when applied to language varieties; (iv) Reusability of language resources in NLP applications (e.g., for machine translation, POS tagging, syntactic parsing, etc.); (v) Corpus-driven studies in dialectology and language variation; (vi) Computational approaches to mutual intelligibility between dialects and similar languages; (vii) Automatic identification of lexical variation; (viii) Automatic classification of language varieties; (ix) Text similarity and adaptation between language varieties; (x) Linguistic issues in the adaptation of language resources and tools (e.g., semantic discrepancies, lexical gaps, false friends); (xi) Machine translation between closely related languages, language varieties and dialects.

<https://sites.google.com/view/vardial-2025>

Time: 09:00-17:30

Location: Conf. Hall B Room A

Sunday-Monday, January 19-20, 2025

Workshop 10: The Joint Workshop of the 9th Financial Technology and Natural Language Processing (FinNLP), the 6th Financial Narrative Processing (FNP), and the 1st Workshop on Large Language Models for Finance and Legal (LLMFinLegal)

Organizers: Chung-Chi Chen, Qianqian Xie, Jimin Huang, Sophia Ananiadou, Hsin-Hsi Chen, Paloma Martínez, Doaa Samy, Mo El-Haj, Paul Rayson, Xiao-Yang (Yanglet) Liu, Benyou Wang, Hao Wang, Jordan Suchow, Alejandro Lopez-Lira, Eghbal Rahimikia, Min Peng, Hen-Hsen Huang, and Hiroya Takamura

The joint workshop of FinNLP, FNP, and LLMFinLegal aims to explore the intersection of Natural Language Processing (NLP), Machine Learning (ML), and Large Language Models (LLMs) within the financial and legal domains. By merging the expertise and focus areas of these three workshops, we intend to foster interdisciplinary research and innovation, addressing the multifaceted challenges inherent in these fields.

<https://sites.google.com/nlg.csie.ntu.edu.tw/finnlp-fnp-llmfinlegal/>

Time: 09:00-17:10

Location: Conf. Hall B Room B

Monday, January 20, 2025

Workshop 11: The Second COLING Workshop on Bridging Neurons and Symbols for NLP and Knowledge Graph Reasoning (BNS-KGR)

Organizers: Kang Liu, Yangqiu Song, Zhen Han, Rafet Sifa, Shizhu He, and Yunfei Long

Recent exploration shows that LLMs, e.g., ChatGPT, may pass the Turing test in human-like chatting but have limited capability even for simple reasoning tasks. It remains unclear whether LLMs reason or not. Human reasoning has been characterized as a dual-process phenomenon or as mechanisms of fast and slow thinking. These findings suggest two directions for exploring neural reasoning: starting from existing neural networks to enhance the reasoning performance with the target of symbolic-level reasoning, and starting from symbolic reasoning to explore its novel neural implementation. These two directions will ideally meet somewhere in the middle and will lead to representations that can act as a bridge for novel neural computing, which qualitatively differs from traditional neural networks, and for novel symbolic computing, which inherits the good features of neural computing. Hence the name of our workshop, with a focus on Natural Language Processing and Knowledge Graph reasoning. This workshop promotes research in both directions, particularly seeking novel proposals from the second direction.

<https://neusymbridge.github.io/>

Time: 08:50-18:00

Location: Conf. Hall B Room A

Workshop 12: 18th Workshop on Building and Using Comparable Corpora (BUCC)

Organizers: Serge Sharoff, Ayla Rigouts Terryn, Pierre Zweigenbaum, and Reinhard Rapp

In the language engineering and linguistics communities, research in comparable corpora has been motivated by two main reasons. In language engineering, on the one hand, it is chiefly motivated by the need to use comparable corpora as training data for statistical NLP applications such as statistical and neural machine translation or cross-lingual retrieval. In linguistics, on the other hand, comparable corpora are of interest because they enable cross-language discoveries and comparisons. It is generally accepted in both communities that comparable corpora consist of documents that are comparable in content and form in various degrees and dimensions across several languages. Parallel corpora are on the one end of this spectrum, and unrelated corpora are on the other. In recent years, the use of comparable corpora for pre-training Large Language Models (LLMs) has led to their impressive multilingual and cross-lingual abilities, which are relevant to a range of applications, including Information Retrieval, Machine Translation, Cross-lingual text classification, etc. The linguistic definitions and observations related to comparable corpora can improve methods to mine such corpora for applications of statistical NLP, for example, to extract parallel corpora from comparable corpora for neural MT or to improve cross-lingual transfer of LLMs. As such, it is of great interest to bring together builders and users of such corpora.

<https://comparable.lisn.upsaclay.fr/bucc2025/>

Time: 09:15-17:45

Location: Capital Suite 9 (Level 2)

Workshop 13: The Fifth Celtic Language Technology Workshop (CLT)

Organizers: Brian Davis, Theodorus Franssen, Elaine Uí Dhonnchadha, Abigail Walsh, and Jane Adkins

The CLTW community and workshop – inaugurated at COLING (Dublin) in 2014 – has become a critical focus and forum for researchers working in natural language processing (NLP) and language technologies for Celtic languages. In particular, it has galvanised and catalysed research by facilitating communication and collaboration internationally. Our community is interested in language technology for both contemporary and historical stages of the Celtic languages.

<https://cltworkshop.github.io/>

Time: 09:00-13:25

Location: Online Virtual Only

Workshop 14: Workshop of Evaluation of Multimodal Generation (EvalMG)

Organizers: Wei Emma Zhang, Xiang Dai, Desmond Elliot, Byron Fang, Haojie Zhuang, Mong Yuan Sim, and Weitong Chen

Multimodal generation techniques have opened new avenues for creative content generation. However, evaluating the quality of multimodal generation remains underexplored and some key questions are unanswered, such as the contributions of each modal, the utility of pre-trained large language models for multimodal generation, and measuring faithfulness and fairness in multimodal outputs. This workshop aims to foster discussions and research efforts by bringing together researchers and practitioners in natural language processing, computer vision, and multimodal AI. Our goal is to establish evaluation methods for multimodal research and advance research efforts in this direction.

<https://evalmg.github.io/>

Time: Check online for Schedule

Location: Capital Suite 10 (Level 2)

Workshop 15: IndoNLP: The First Workshop on Natural Language Processing for Indo-Aryan and Dravidian Languages

Organizers: Ruwan Weerasinghe, Isuri Anuradha, Deshan Sumanathilaka, Mo El-Haj, Chamila Liyanage, Fahad Khan, Andrew Hardie, Asim Abbas, Ruslan Mitkov, Julian Hough, Nicholas Micallef, and Naomi Krishnarajah

The rapid advancement of Natural Language Processing (NLP) and Large Language Models (LLMs) has transformed the landscape of computational linguistics. However, Indo-Aryan and Dravidian Languages (IADL), which represent a significant portion of South Asia's linguistic heritage, remain under-resourced and under-researched in these technological developments. This workshop aims to bridge this gap by bringing together researchers, linguists, and technologists to focus on the unique challenges and opportunities. Participants will explore innovative methods for creating and annotating digital corpora, develop speech and language technologies suited to IADL, and promote interdisciplinary collaborations. By leveraging LLMs, we seek to address the complexities of syntax, morphology, and semantics in these languages to enhance the performance of NLP applications. Furthermore, the workshop will provide a platform for sharing best practices, tools, and resources, enhancing the digital infrastructure necessary for language preservation. Through collaborative efforts, we aim to build a research community to advance NLP for IADL, contributing to linguistic diversity and cultural preservation in the digital age.

<https://indonlp-workshop.github.io/IndoNLP-Workshop/>

Time: 08:45-17:00

Location: Conf. Hall B Room D

Workshop 16: The First Workshop on Language Models for Low-Resource Languages (LoResLM 2025)

Organizers: Hansi Hettiarachchi, Tharindu Ranasinghe, Paul Rayson, Ruslan Mitkov, Mohamed Gaber, Damith Premasiri, Fiona Anting Tan, and Lasitha Uyagodage

Neural language models have revolutionised natural language processing (NLP) and have provided state-of-the-art results for many tasks. However, their effectiveness is largely dependent on the pre-training resources. Therefore, language models (LMs) often struggle with low-resource languages in both training and evaluation. Recently, there has been a growing trend in developing and adopting LMs for low-resource

languages. This workshop aims to provide a forum for researchers to share and discuss their ongoing work on LMs for low-resource languages.

<https://loreslm.github.io/>

Time: 08:45-18:00

Location: Capital Suite 5 (Level 2)

Workshop 17: The 1st International Workshop on Nakba Narratives as Language Resources (Nakba-NLP 2025)

Organizers: Mustafa Jarrar, Nizar Habash, Mo El-Haj, Amal Haddad, Zeina Jallad, Camille Mansour, Diana Allan, Paul Rayson, and Tymaa Hammouda

The narratives of the (ongoing) Palestinian Nakba possess significant historical, cultural, literary, and academic value. Preserving this content and empowering it with AI tools is crucial for ensuring its accessibility and usability for present and future generations. Nakba narratives and testimonies exist in diverse formats such as manuscripts, books, audio recordings, novels, and films. Converting this content into a machine-understandable format presents a notable challenge. Establishing accessible archives and well-annotated collections is essential for researchers and historians to verify and share meaningful information. This workshop aims to explore how artificial intelligence, natural language processing, and corpus linguistics can assist in understanding, disseminating and preserving, Nakba narratives and testimonies. The goal is to create accessible, comprehensive, and well-annotated collections that empower researchers and historians to validate and share critical insights derived from these data. The workshop targets datasets and narratives in Arabic, English, and other languages, however, submitted articles should be written in English.

<https://sina.birzeit.edu/nakba-nlp/>

Time: 09:00-17:15

Location: Online Virtual Only

Workshop 18: RegNLP-2025: The First Workshop on Regulatory NLP

Organizers: Annie Antón, Barry West, Tuba Gokhan, Kexin Wang, Iryna Gurevych, and Ted Briscoe

The complexity, volume, and ever-changing nature of regulatory documents present unique challenges in governance, compliance, and legal frameworks across various sectors. Addressing these challenges demands specialized approaches in natural language processing (NLP) to enable effective management and utilization of regulatory content. Recent advancements in NLP have opened new avenues for tackling these issues, specifically tailored to the domain of regulatory documents. These advancements include sophisticated techniques for document parsing, entity recognition, and automated compliance checking, which are essential for navigating the intricate landscape of regulatory requirements. Despite these technological strides, significant open questions and challenges remain. How can NLP models better handle the dynamic and diverse nature of regulatory texts? What methods are most effective for extracting and synthesizing information from vast and complex document repositories? How can we ensure the accuracy and reliability of automated compliance tools? Moreover, what are the best practices for adapting NLP models to the highly specialized language and context of regulatory documents? The first workshop on Regulatory Natural Language Processing (RegNLP) aims to convene a diverse group of researchers and practitioners from NLP, legal informatics, compliance, and related fields to explore these questions.

<https://regnlp.github.io/>

Time: 09:00-17:30

Location: Capital Suite 13 (Level 2)

Workshop 19: The Second Workshop on South East Asian Language Processing (SEALP)

Organizers: Derry Wijaya, Alham Fikri Aji, Clara Vania, Genta Indra Winata, Ayu Purwarianti

South East Asia is one of the most linguistically diverse regions in the world, with over 1200 languages spoken by 680 million people. However, the diversity of South East Asian languages has long been at risk due to the emphasis on national languages as lingua franca in South East Asian countries at the end of colonization; and the increasing prominence of English due to the necessities of globalization. This workshop will bring together practitioners from academia, government, and industry interested in the research and development of language technologies for SEA languages. The workshop also aims to build an inclusive community of everyone passionate about SEA languages, increase community awareness of works that have been developed to date on these languages, and foster collaborations that will strengthen and spur NLP research and development in SEA languages.

<https://sealp-workshop.github.io/>

Time: Check online for Schedule

Location: Online Virtual Only

Workshop 20: SUMEval-2: The 2nd Workshop on Scaling Up Multilingual & Multi-Cultural Evaluation

Organizers: Hellina Hailu Nigatu, Monojit Choudhury, Oana Ignat, Sebastian Ruder, Sunayana Sitaram, and Vishrav Chaudhary

Massively Multilingual Language Models (MMLMs) like mBERT, XLMR and XY-LENT support around 100 languages of the world. Additionally, generative models like GPT-4 and BLOOM are getting attention from the NLP community and the public. However, most existing multilingual NLP benchmarks reflect a handful of cultures and languages. The languages present in evaluation benchmarks are usually high-resource and largely belong to the Indo-European language family. By extension, the cultures represented in evaluation benchmarks are also largely reflective of Western society. This makes current evaluation unreliable and does not provide a full picture of the performance of MMLMs across the linguistic and cultural landscape. Although efforts are being made to create benchmarks that cover a larger variety of tasks, cultures, languages, and language families, it is unlikely that we will be able to build benchmarks covering all languages and cultures. Due to this, there is recent interest in alternate strategies for evaluating MMLMs, including performance prediction and Machine Translation of test data.

<https://sites.google.com/view/sumeval-2025>

Time: 09:00-17:30

Location: Capital Suite 6 (Level 2)

Workshop 21: The 4th Workshop on Arabic Corpus Linguistics (WACL-4)

Organizers: Saad Ezzini, Hamza Alami, Ismail Berrada, Abdessamad Benlahbib, Abdelkader El Mahdaouy, Salima Lamsiyah, Hatim Derrouz, Amal Haddad, Mustafa Jarrar, Mo El-Haj, Ruslan Mitkov, and Paul Rayson

The field of Arabic language research using corpora and corpus methods has experienced significant growth and development in recent years. What once were isolated efforts have now transformed into a vibrant and expansive area of study, advancing rapidly across multiple dimensions in both corpus and computational linguistics. Building upon the success of previous editions—WACL-1 in 2011, WACL-2 in 2013 in conjunction with the Corpus Linguistics Conference at Lancaster University, and WACL-3 in 2019 at the Corpus Linguistics 2019 conference at Cardiff University—we are excited to announce the fourth edition of the Workshop on Arabic Corpus Linguistics (WACL-4). The primary objectives of

WACL-4 are to highlight the latest developments in the creation, annotation, and application of Arabic corpora, including the introduction of new corpora and advancements in annotation techniques, while fostering collaboration among researchers from diverse institutions and regions to stimulate joint research projects and interdisciplinary initiatives. This edition will place a special emphasis on the study of Arabic dialects, including non-standard and regional varieties, to broaden the understanding of Arabic in its various manifestations and support research on under-resourced linguistic varieties. Additionally, WACL-4 aims to encourage the development and refinement of Natural Language Processing (NLP) systems and tools tailored for Arabic, integrating corpora into NLP workflows, creating new computational tools, and evaluating existing systems to improve their efficacy in processing Arabic text.

<https://wp.lancs.ac.uk/wacl4/>

Time: 09:00-16:30

Location: Online Virtual Only

Workshop 22: Writing Aids at the Crossroads of AI, Cognitive Science and NLP WR-AI-CogS

Organizers: Michael ZOCK, Kentaro Inui, and Zheng Yuan

This workshop is dedicated to developing writing aids that align with human cognition, including attention and memory limitations, common writing habits, knowledge states, and information needs. In other words, we focus on the cognitive and engineering aspects of interactive writing. Our goal is not only to help people acquire and improve their writing skills but also to enhance their productivity. By leveraging computer technology, we aim to enable them to produce better texts in less time.

<https://sites.google.com/view/wraicogs1>

Time: 09:00-18:30

Location: Capital Suite 8 (Level 2)

Conference Venue



COLING 2025 is taking place in the Abu Dhabi National Exhibition Centre (ADNEC). ADNEC is a multi-award-winning venue offering organizers of exhibitions, conferences, and events outstanding facilities spread over a total space of 151,000 square meters. ADNEC is just 15 minutes away from the rapidly expanding Abu Dhabi International Airport which serves over 50 airlines and 100 destinations in 46 countries around the globe. It is only 20 minutes from downtown Abu Dhabi and 90 minutes from Dubai International Airport.

- Address: Abu Dhabi National Exhibitions Company, Khaleej Al Arabi Street, P.O. Box 5546, Abu Dhabi, United Arab Emirates
- Phone: +971 (0) 2 444 6900

- Web: <https://adnec.ae/>
- E-mail: customer.services@adnec.ae

Abu Dhabi



Abu Dhabi, the capital of the United Arab Emirates, is a city of contrasts where traditional meets modern, offering visitors an unforgettable experience. Situated on a set of islands in the Arabian Gulf, Abu Dhabi is a cosmopolitan city with a very large international presence (90% of residents). English is the lingua franca.

Local Transportation in Abu Dhabi

- ADNEC is just a 20-minute car ride from Abu Dhabi Airport. From Abu Dhabi Corniche, ADNEC is located just 15 minutes away.
- Taxis are available throughout Abu Dhabi and they are quite convenient, safe and affordable. You can order a taxi by calling 971 600 53 53 53 or using the Yango app (AppStore, Google Play). A number of taxis (all reliable) are available at Abu Dhabi airport, and from ADNEC at the adjacent Aloft Hotel entrance. UBER is also available in Abu Dhabi but it is generally a bit more expensive than taxis.
- Buses are also available around Abu Dhabi and ADNEC is serviced by bus number 040, seven days a week. More information about bus routes and timings can be found on Google Maps (AppStore, Google Play). A weekly subscription for public transport costs 30 AED and can be purchased at Abu Dhabi Airport, vending machines in the city, and Lulu Supermarkets.

Childcare

Please check with the specific hotel where you plan to stay for childcare support.

Dining Options

There are a variety of dining selections located near ADNEC. A short walk or drive away, chances are you

will find something to satisfy your tastebuds. Use a search engine to find a restaurant near you. There are also a number of cafes and two supermarkets (Zoom and Lulu Express) within a 5 minute walk from ADNEC. Below are recommendations made by a local chair:

- **Saudi Kitchen** <https://maps.app.goo.gl/JXDFdY5UxesJ6Wcu8>
 - Attractive atmosphere (floor seating)
 - Excellent mandi
 - Suitable for large groups (10+)
 - Very affordable
 - 8-minute ride from ADNEC
- **Mosaic** <https://maps.app.goo.gl/gkCfGz8xyDaX4yeS8>
 - Elegant but down to earth
 - Excellent Lebanese dishes
 - Suitable for group of 8-10
 - Affordable
 - 8-minute ride from ADNEC
- **Bait El Khetyar** <https://maps.app.goo.gl/x8aehmUH2Bh4rY3m8>
 - Casual dining
 - Excellent hummus, mutabal, and shawarma
 - Limited seating (but fast for take-away)
 - Very affordable
 - 8-minute drive from the Corniche
- **Tasha**
 - Casual dining in a nice atmosphere
 - Excellent Egyptian dishes (mahshi, koshari, etc.)
 - Good for groups of 8-10
 - Affordable
 - 2-minute ride from the Corniche

Abu Dhabi Landmarks

Emirates Palace

Located in the heart of Abu Dhabi, Emirates Palace is one of the capital's most well-known landmarks. It is known for its enchanting Arabesque style, award-winning five-star hospitality, and wonderfully unique and authentic local experiences.

Louvre Abu Dhabi

The iconic Louvre Abu Dhabi is the first universal museum in the Arab World, translating and fostering the spirit of openness between cultures. As one of the premier cultural institutions located in the heart of the Saadiyat Cultural District on Saadiyat Island, this art-lovers' dream displays works of historical, cultural, and sociological significance, from ancient times to the contemporary era.

Sheikh Zayed Grand Mosque

The impressive and inspiring Sheikh Zayed Grand Mosque is one of the world's largest mosques and the only one that captures the unique interactions between Islam and other world cultures. Sheikh Zayed bin Sultan Al Nahyan, the Founder of the UAE, had a very specific vision for this mosque: to incorporate architectural styles from different Muslim civilizations and celebrate cultural diversity by creating a haven that is truly welcoming and inspirational in its foundation. The mosque's architects were British, Italian, and Emirati, with design ideas borrowed from parts of Turkey, Morocco, Pakistan, and Egypt, among other Islamic countries.

Qasr Al Hosn

Over the centuries, Qasr Al Hosn has been home to the ruling family, acted as the seat of government, housed the National Consultative Council founded by the late Sheikh Zayed Bin Sultan Al Nahyan, Founder of the UAE, as well as being a national archive. Today it stands as the nation's living memorial and a narrator of Abu Dhabi's history.

The National Aquarium Abu Dhabi

The largest aquarium in the Middle East, The National Aquarium in Al Qana is literally swimming with aquatic wildlife featuring over 46,000 animals from more than 300 unique species. Spread across 10 nautically-themed zones, from the UAE's Natural Treasures, sunken sea wrecks, and Atlantic caves, right through to flooded forests, fiery volcanoes, and a frozen ocean, there are more than 60 attractions that will be sure to delight and excite the whole family.

Yas Island

This family-friendly entertainment hub is just a 30-minute drive from the city and a 15-minute drive from Abu Dhabi International Airport. Yas Island is home to an F1™ race track, several theme parks, an incredible links golf course, a beautiful beach, stunning hotels, an impressive mall and much more. It's the perfect place to bring the kids, from toddler to teen, for a fun-filled holiday by the sea.

Jubail Mangrove Park

Jubail Mangrove Park is home to meandering boardwalks that allow you to wander through the mangroves, discovering a nature-rich side to Abu Dhabi and spotting an array of wildlife, from turtles to herons, gazelle, and more. Birdwatchers, nature-lovers, and photographers will be in their element here.

Abu Dhabi's Corniche

Abu Dhabi's Corniche stretches over eight kilometers of manicured waterfront that includes children's play areas, separate cycle and pedestrian pathways, cafés and restaurants, and a lifeguarded beach.

Abu Dhabi Beaches

- **Soul Beach Saadiyat Island:** Tucked away in Saadiyat Island, Soul Beach is Saadiyat Island's new captivating beachfront location, part of the Mamsha Al Saadiyat community.
- **Saadiyat Beach:** With pristine white sands stretching out along a generous shore, Saadiyat Beach maintains its reputation as one of the most desirable beach locations in the UAE.
- **Yas Beach:** Set on a majestic stretch of white sand, sun-seekers holidaying at any of Yas Island's hotels and hotel apartments can enjoy complimentary access to the island's crystal clear waters and natural mangrove surrounds.
- **Corniche Beach:** This beautiful beach isn't just home to turquoise water and soft, white sand, it also has a beautiful seaside boardwalk, with well-kept walkways home to manicured gardens and benches overlooking the picturesque Arabian Gulf.
- You might notice a location not far from ADNEC called "**Al Qurm Beach**". While this is not a swimming spot, it is a lovely area for a walk, where many locals gather for BBQs and games.

Important Information

Climate

Abu Dhabi has a northern-hemisphere subtropical, arid climate. November to March is the most appealing time of year, and it is also when infrequent winter rains occur. The temperature range in the winter months is between 56 F (13 C) and 75 F (24 C) with typically bright sunny days which correspond to the best kind of spring weather in the US.

Clothing

Summer clothing may be worn for most of the year, but during the winter evening temperatures may occasionally call for a jacket or light coat. While dress codes are fairly liberal, consideration should be given not to offend the sensibilities of others. Swimwear should be worn only on beaches or at swimming pools. Visiting shopping malls and other attractions, tourists should wear clothing that is not too tight or revealing. Certain attractions such as mosques or religious sites usually have stricter dress codes, requiring both men and women to cover up bare shoulders, arms and legs, and women to wear headscarves.

Communications

The international dialing code for incoming calls to landlines in the UAE is +971 and 02 for Abu Dhabi. Calls to and from landlines within Abu Dhabi are free. Direct dialing is possible to over 170 countries. UAE has two mobile networks, Du and Etisalat:

- Both offer temporary SIM cards for tourists and business travelers, including data and calls, and these can be purchased at outlets across the UAE, including at the airport and malls.
- Roaming services are also available for most visitors if they wish to use their existing number and phone.

Currency and Living Cost

The monetary unit is the Dirham (AED), which is divided into 100 fils. The exchange rate is pegged to the US Dollar at the rate \$1 = AED 3.675.

The average costs at an average coffee shop or restaurant are as follows:

- Cup of coffee: 14-20 AED/\$3.8-5.5 USD
- Sandwich lunch: 15-30 AED/\$5-8.2 USD
- Evening meal: 70-100 AED/\$19-27.2 USD

Electricity

The electricity supply in Abu Dhabi is 220/240 volts at 50 cycles. Standard British-type 13-amp square

three-pin plugs are the norm in most hotels. European or US-made appliances may need a plug adapter.

Language

The official language of the UAE is Arabic. However, English is also very widely spoken throughout Abu Dhabi, especially in business, hospitality, retail environments, street signs, taxis, restaurant menus, etc. Urdu and Hindi are also widely spoken.

Top eight phrases to learn:

- Marhaba - Hello
- Ahlan - Welcome
- Ma'asalama - Good Bye
- Shukran - Thank you
- Mabrook - Congratulations
- Yalla - Let's go
- Khalas - enough/done
- Inshallah - God willing, may be, no, and a host of other context dependent meanings.

Security

As one of the most cosmopolitan and multicultural cities in the world, home to over 200 different nationalities, Abu Dhabi is an advocate for peace and stability, and proud to be a connecting hub between East and West. Abu Dhabi is ranked in the 10 safest cities by Aon Hewitt, with low crime rates, a stable government, and a department of Abu Dhabi Police dedicated entirely to visitors.

Taxation

The UAE does not levy income tax on individuals. Value Added Tax is levied on a majority of goods and services.

Time Zone & Business Hours

Abu Dhabi is GMT+4. Most businesses are open from 8 am to 6 pm, Monday to Friday, with Saturday and Sunday being official holidays for all government departments. Embassies, consulates, and government offices operate from 7:30 am to 2:30 pm, Monday to Friday.

Tipping

Tipping practices are similar to most other parts of the world. Most restaurants include a 10% service charge, but tipping, in general, is at the customer's discretion.

Water

The tap water in Abu Dhabi is safe to drink. But locally bottled water is generally served in hotels and restaurants.

Emergency

The emergency phone number for Abu Dhabi Police is 999. Whether you need police assistance, an ambulance, or for any other emergency, 999 is the number to call and calls are free. When calling 999, please remember to state your name, the nature of the accident, address of the emergency and how serious the situation is. Please find a detailed list of other emergency numbers at the bottom of this page.

If you're involved in a traffic accident, it's important to contact the police immediately. In case of a minor incident, move your car to the roadside, as there are fines for obstructing traffic. Remember, you cannot file an insurance claim without a police report.

For other enquiries, Abu Dhabi Police operates a dedicated Tourism Police section which will advise and guide you on a range of matters. You can contact them on +971 2 800 2626 and +971 2 512 7777, or visit Emergency Help for Tourists.

In a medical emergency, Abu Dhabi's Sheikh Khalifa Medical City (+971 2 819 0000) and Mediclinic Al Noor Hospital (800 2000) both have Accident and Emergency units. If you're injured in a traffic accident,

you will automatically be taken to Sheikh Khalifa Medical City, as it has the best Accident & Emergency treatment facilities.

The Abu Dhabi Government portal provides an updated list of 24-hour pharmacies and medical services, including hospitals, clinics, and medical centres. If you don't have internet access you can call the toll-free number 800 555 (+971 2 666 4442).

Dubai



Dubai is only 140km (1 hour and 20 minutes by car) away from Abu Dhabi. Dubai is a vibrant, urban, and multicultural city with attractive tourist destinations, such as Burj Khalifa – the tallest building in the world.

Sharjah



Sharjah is the third largest Emirate, it is only 165km (1 hour and 50 minutes by car) away from Abu Dhabi. It is a major cultural hub in the UAE, there is also a wide variety of activities awaiting.

14

Venue Map

The Convention Centre

COLING 2025 is held in the Abu Dhabi National Exhibition Centre (ADNEC).

The Registration Desk (January 18 to January 24, 2025) is in The Link (next to Hall B).

Workshops and tutorials (January 19 & 20, 2025) are in Capital Suites 5 to 10, 13. Halls B-A, B-B, B-C, B-D.

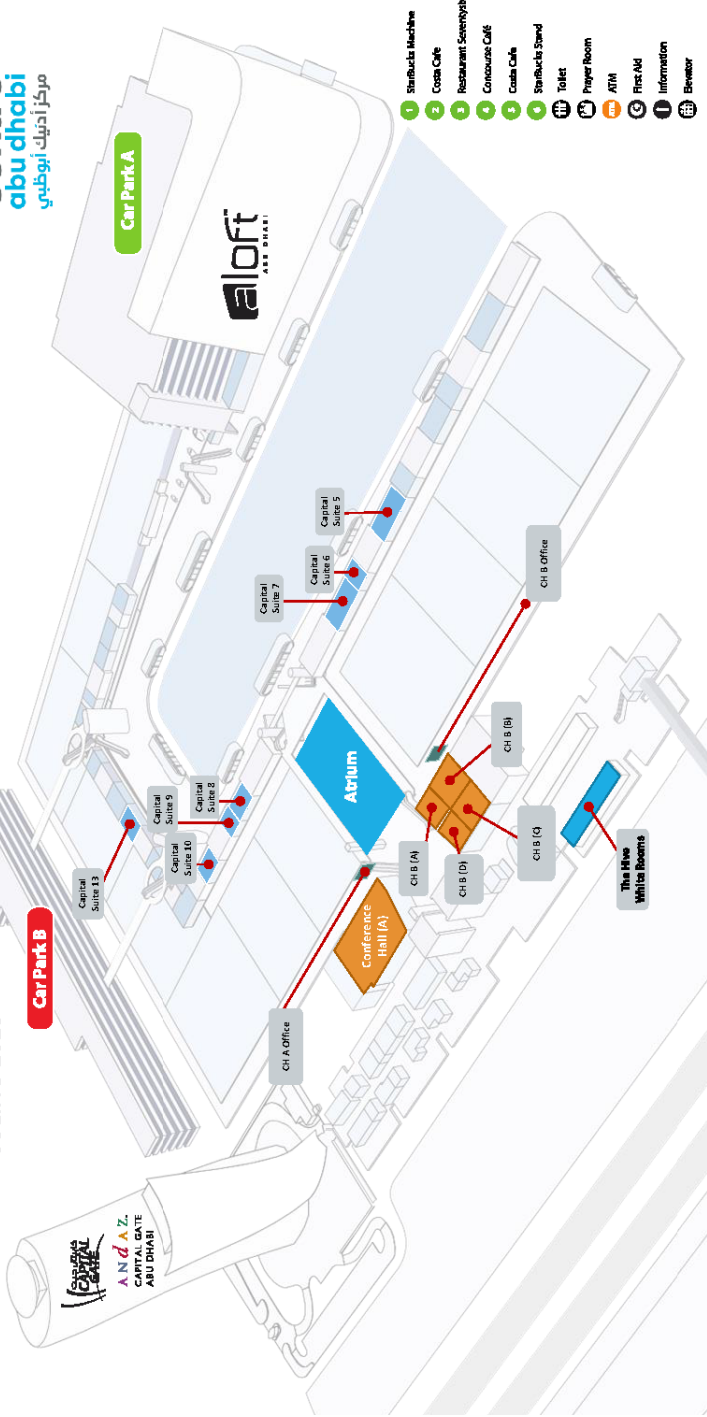
Coffee breaks for the Main Conference are held in the Link and Atrium. Coffee breaks for Workshop are held in the walkways outside the capital suites and Hall B.

The Main Conference (January 21 to 24, 2025) is held in Hall A (Plenary), Hall B Rooms A-D, Capital Suite 7 (Oral Group Discussions). Group Discussions for Demos, posters, workshops along with exhibitions, will be located in The Link and The Atrium areas, between and adjacent to Halls A and B.

VENUE MAP

International Conference on Computational Linguistics COLING 2025

adnec
centre
abu dhabi
مركز ادنيك ابوظبي





MOHAMED BIN ZAYED
UNIVERSITY OF
ARTIFICIAL INTELLIGENCE

POWER from
KNOWLEDGE to SERVE

Join the World's First Graduate AI University



Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), the world's first graduate AI university, is a research-focused institution with **199 M.Sc.** and **87 Ph.D.** students from **39 nationalities (30% female)** and **81 renowned faculty members**. Ranked **8th in NLP** and **10th in AI+NLP+CV+ML** (csranks.org), MBZUAI's NLP department comprises **19 faculty** and **80 graduate students**.

Now hiring **Full, Associate, and Assistant Professors**, along with **graduate student opportunities** for M.Sc. or Ph.D. programs.

Faculty & Postdoctoral Positions in NLP

We're recruiting:

- ▶ Full Professor
- ▶ Associate Professor
- ▶ Assistant Professor
- ▶ Postdoctoral Researchers

Graduate Student Opportunities

Pursue your M.Sc. or Ph.D. at MBZUAI:

- ▶ Access to **world-class faculty** and **research infrastructure**
- ▶ Focused programs in **Natural Language Processing (NLP)**



Faculty & Postdoctoral Positions in NLP. For more information on M.Sc. and Ph.D. application requirements, programs, and how to apply, scan the QR code



mbzuai.ac.ae

MBZUAI



CREATING NEW OPPORTUNITIES FOR YOUR TOMORROW

At the Technology Innovation Institute, we push the boundaries of applied research with a focus on the real world of tomorrow. We make a difference to the future of the UAE and our planet by working at scale and at speed across ten dedicated research centers. We cultivate excellence and curiosity, not simply asking "why?" but also "why not?" and "what if?".

From state-of-the-art facilities in Abu Dhabi to global partnerships, we catalyze change in the fields of AI, renewable energy, advanced materials, autonomous robotics, biotechnology, cryptography, directed energy, propulsion and space technologies, secure systems, and quantum engineering.

Learn more about us on tii.ae



NYU Abu Dhabi welcomes and educates global citizens, and produces knowledge that promotes human understanding and betters society.

Through distinctive global liberal arts education and graduate programs, we enable students and graduates to achieve intellectual, personal, and professional fulfillment and empower them to make significant societal contributions to Abu Dhabi and the world. Through cutting-edge research, we develop knowledge, foster creativity and innovation, and help solve humanity's shared challenges. Together, we contribute to Abu Dhabi's knowledge-based economy and society, and play a central role in NYU's global mission.

جامعة نيويورك ابوظبي



NYU ABU DHABI



JAIS **Family of Models**

JAIS is a family of up to 70-billion parameter LLM's, delivering superior performance and efficiency in tasks such as summarization, comprehension, and text generation.



Equitable access to AI

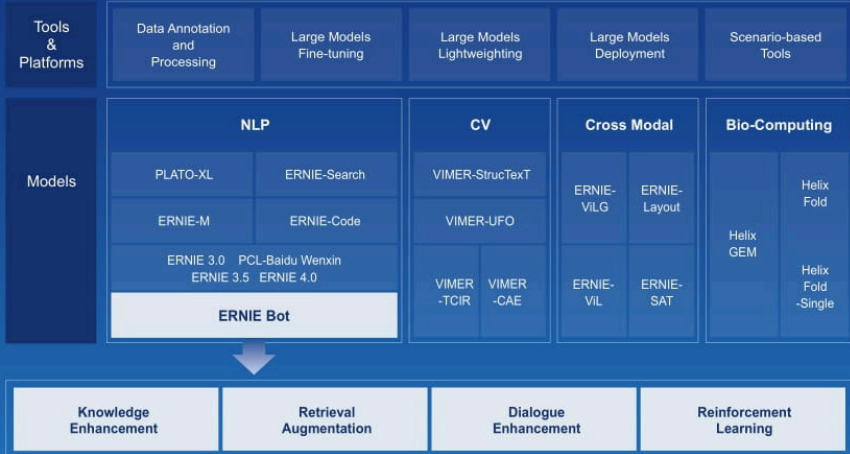
-  Fully bilingual
-  Bringing AI to everyone
-  Optimized for Arabic adaptive fine-tuning
-  Easily Integrated wherever you build

Join us at the event between January 19th to 24th and learn how AI models can drive efficiency, unlock insights, and enhance your decision-making.

BAIDU WENXIN

A series of knowledge-enhanced large models, from general-purpose to industry-specific, independently developed by Baidu, speed up AI innovations and empower the industry to upgrade.

ERNIE Knowledge - Enhanced Large Models



Your high quality data partner for GenAI development

Demonstrations collection for SFT

Preferences collection for RLHF / DPO

Evaluation and Red Teaming

40+
languages supported

10+
code languages

20+
knowledge domains

10k
prompt-completion pairs per week per domain

200k
aligned pairs per week

Visual, audio and creative models

Responsible AI

Autonomous agents

Text generation models

Coding assistants



toloka.ai



COLING 2025

SPONSORS

**DESTINATION
PARTNER**

abu dhabi
Convention & Exhibition Bureau



**HOST
INSTITUTION**

**DIAMOND
SPONSOR**



جامعة نيويورك أبوظبي
 NYU ABU DHABI

**ARABIC NLP
WINTER SCHOOL
SPONSOR**

**PLATINUM
SPONSOR**



**GOLD
SPONSOR**

**SILVER
SPONSORS**



**BRONZE
SPONSOR**

**SUPPORTER
SPONSOR**

