1st Call for Papers

The 3rd Annual Meeting of the ELRA-ISCA Special Interest Group on Under-resourced Languages (SIGUL2024)

A Satellite Workshop of [LREC-COLING 2024](#)

Monday and Tuesday, May 20th-21st, 2024
Torino, Italy (co-located with LREC-COLING 2024)
Workshop website: https://sigul-2024.ilc.cnr.it (under construction)

- Deadline for submissions: February 26th, 2024
- Paper submission link: https://softconf.com/lrec-coling2024/sigul2024/
- Deadline for camera-ready papers: April 5th, 2024

The 3rd Annual Meeting of the [ELRA](#)/[ISCA](#) [Special Interest Group on Under-Resourced Languages](#) (SIGUL2024) will provide a forum for the presentation and discussion of cutting-edge research in language processing for under-resourced languages by academic and industry researchers. Following the long-standing series of previous meetings, the SIGUL workshop will also offer a venue where researchers in different disciplines and from varied backgrounds can fruitfully explore new areas of intellectual and practical development while honoring their common interest of sustaining less-resourced languages.

Topics

We invite contributions (regular long papers of 8 pages or short papers of 4 pages) targeting any of the following - non-exhaustive - list of topics:

- Processing any under-resourced languages (covering less-resourced, under-resourced, endangered, minority, and minoritized languages)
- Cognitive and linguistic studies of under-resourced languages
- Fast resources acquisition: text and speech corpora, parallel texts, dictionaries, grammars, and language models
- Zero and few-shot methodologies and self-supervised learning in language and speech technologies
- Cross-lingual and multilingual acoustic and lexical modeling
- Speech recognition and synthesis for under-resourced languages and dialects
- Machine translation and speech-to-speech translation
- Spoken dialogue systems
- Applications of language technologies for under-resourced languages
- Large language models and under-resourced languages

Special Topic

- Text and speech resources and technologies for the languages of Italy

Special Session on languages of Italy and language technologies

Italy is known for its linguistic diversity that reflects its long and varied history. To celebrate it, SIGUL2024 will provide a special session or forum for researchers interested in developing language resources and technologies for the many languages of Italy (regional, minority, or heritage languages, including those of the neighboring countries).

Submissions

Authors can choose among three paper categories:

- Regular long papers – up to eight (8) pages maximum*, presenting substantial, original, completed, and unpublished work.

- Short papers – up to four (4) pages*, describing work-in-progress projects in the early stage of development, new resources, negative results, system demonstrations, and early-career/student work.
- Position papers – up to eight (8) pages*, for reflective considerations of methodological, best practice, and institutional issues (e.g., ethics, data ownership, speakers' community involvement, de-colonizing approaches).

The above page limits exclude any number of additional pages that may be needed for references.

The form of the presentation may be oral or poster, whereas in the proceedings there is no difference between the accepted papers. Submission is NOT anonymous, and the official LREC-COLING 2024 format must be adopted. Each paper will be reviewed by three independent reviewers.

Invited speakers

TBA

Important Dates

- 26 February 2024: submission due
- 18 March 2024: reviews due
- 22 March 2024: notifications to authors
- 5 April 2024: camera-ready (PDF) due

Identify, Describe and Share your LRs!

When submitting a paper from the START page, authors will be asked to provide essential information about resources (in a broad sense, i.e. also technologies, standards, evaluation kits, etc.) that have been used for the work described in the paper or are a new result of your research. Moreover, ELRA encourages all LREC-COLING authors to share the described LRs (data, tools, services, etc.) to enable their reuse and replicability of experiments (including evaluation ones).

Workshop Organizers

Maite Melero, Sakriani Sakti, Claudia Soria

Program Committee

- Mohammad A. M. Abushariah (The University of Jordan, Jordan)
- Manex Agirrezabal (University of Copenhagen – Center for Sprogteknologi | Center for Language Technology, Denmark)
- Shyam S. Agrawal (KIIT, Gurugram, India)
- Begoña Altuna (HiTZ Center - Ixa, Euskal Herriko Unibertsitatea | University of the Basque Country, Spain)
- Antti Arppe (University of Alberta, Canada)
- Martin Benjamin (Kamusi Project International)
- Delphine Bernhard (Université de Strasbourg, LiLPa, France)
- Steven Bird (Charles Darwin University, Australia)
- Claudia Borg (University of Malta)
- Matt Coler (University of Groningen, Campus Fryslân, The Netherlands)
- Dan Cristea (Romanian Academy, Romania)
- Pradip Kumar Das (IIT Guwahati, India)
- Seza Doğruöz (Universiteit Gent, België | Ghent University, Belgium)
- Stefano Ghazzali (Language Technologies Unit Bangor University
- Prifysgol Bangor | Bangor University, Bangor, Gwynedd)
- Itziar Gonzalez-Dios (HiTZ Basque Center for Language Technologies - Ixa, University of the Basque Country UPV/EHU)

- Lars Hellan (Norwegian University of Science and Technology, Norway)
- Mélanie Jouitteau (IKER, CNRS, France)
- Richard Littauer (unaffiliated)
- Teresa Lynn (Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates)
- Nina Markl (University of Essex, UK)
- Maite Melero (Barcelona Supercomputing Center, Espanya | Spain)
- Peter Mihajlik (Budapest University of Technology and Economics, Hungary)
- Win Pa Pa (UCS Yangon, Myanmar)
- Sandy Ritchie (Google Research)
- Sakriani Sakti (JAIST, Japan)
- Claudia Soria (CNR-ILC, Italia | Italy)
- Daan Van Esch (Google Research)
- Menno van Zaanen (South African Centre for Digital Language   Resources, South Africa)
- Jenifer Vega Rodriguez (GIPSA-lab, Université Grenoble Alpes, France)
- Marcely Zanon Boito (NAVER Labs Europe, France)

Contact

mailto:mclaudia.soria@ilc.cnr.it

Please, write "SIGUL2024" in the subject of your e-mail.