



[SIGUL 2024 Workshop](#)

to be held as part of the *Joint International Conference on Computational Linguistics, Language Resources and Evaluation* ([LREC-COLING 2024](#))

Torino, Italy, 20-21 May 2024

2nd Call for Papers

The 3rd Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages (SIGUL 2024) provides 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. After being organized as an Interspeech Workshop in 2023, the SIGUL Workshop Series returns under the umbrella of LREC, which will be held jointly with COLING at the [Lingotto Conference Centre](#) in Torino as well as in a virtual format from 20 to 25 May 2024. As usual, this Workshop will span the research interest areas of less-resourced, under-resourced, endangered, minority, and minoritized languages.

Special Feature

This year, the Workshop will be marked with the following special event:

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, SIGUL 2024 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).

Invited Speakers

Eddie Avila (GlobalVoices)

Jean Maillard (FAIR, META)

Workshop 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 of 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.

Submission

Authors can choose among three paper categories:

- *regular long papers* (up to 8 pages) presenting substantial, original, completed, and unpublished work;
- *short papers* (up to 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 8 pages) for reflective considerations of methodological, best practice and institutional issues (e.g. ethics, data ownership, speakers' community involvement and decolonizing approaches).

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

The form of presentation may be oral or poster, whereas in the Workshop Proceedings there will be no difference among 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.

LREC-COLING 2024 Stylesheet: <https://lrec-coling-2024.org/authors-kit>

SIGUL 2024 Submission Page: <https://softconf.com/lrec-coling2024/SIGUL>

Important Dates

- Paper submission deadline: 26 February 2024
- Notification of acceptance: 22 March 2024
- Camera-ready paper: 5 April 2024
- Workshop dates: 20-21 May 2024

Outline of the Programme

SIGUL 2024 will continue the tradition of the previous SIGUL events featuring a number of distinguished keynote speakers, technical oral and poster sessions and panel discussions to discuss a better future for under-resourced languages and under-resourced communities.

Diversity and Inclusion Subsidies

SIGUL 2024 is providing funds for registration and travel or for bandwidth/VPN. We encourage citizens of developing countries and members of marginalized communities to apply for subsidies. Details on the application procedure will be available on the workshop Web site. For inquiries, please contact claudia.soria@ilc.cnr.it.

Full List of Organizers

SIGUL 2024 Organizing Committee

Maite Melero (BSC, Spain)

Sakriani Sakti (JAIST, Japan)

Claudia Soria (CNR-ILC, Italy)

SIGUL 2024 Programme 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)

A. 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, Spain)

Lars Hellan (Norwegian University of Science and Technology, Norway)

Mélanie Jouitteau (IKER, CNRS, France)

Ritesh Kumar (UnReal-TecE LLP, India)

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)

Nay San (Stanford University, USA)

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)

Marceley Zanon Boito (NAVER Labs Europe, France)

Acknowledgments

SIGUL is a joint Special Interest Group of the [ELRA Language Resources Association \(ELRA\)](#) and of the [International Speech Communication Association \(ISCA\)](#). The SIGUL 2024 Workshop has been organized with the help of the local organizers of [LREC-COLING 2024](#).

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 2024 authors to share the described LRs (data, tools, services etc.) to enable their reuse and replicability of experiments (including evaluation ones).

Contact

To contact the organizers, please send an e-mail to info@sigul.eu (Subject: [SIGUL 2024]).