



## [SIGUL 2023 Workshop](#)

co-located with [Interspeech 2023](#)

Dublin, Ireland, 18-20 August 2023

### **Call for Papers**

The 2<sup>nd</sup> Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages (SIGUL 2023) provides a forum for the presentation and discussion of cutting-edge research in text and speech processing for under-resourced languages by academic and industry researchers. SIGUL 2023 carries on the tradition of the SIGUL and CCURL-SLTU (Collaboration and Computing for Under-Resourced Languages – Spoken Language Technologies for Under-resourced languages) Workshop Series, which have been organized since 2008 and, as LREC Workshops, since 2014. As usual, this Workshop will span the research interest areas of less-resourced, under-resourced, endangered, minority, and minoritized languages.

### **Special Features**

This year, the Workshop will be marked with three special events.

(1) **Special Session on Celtic Language Technology** (August 18)

SIGUL 2023 will provide a Special Session or Forum for researchers interested in developing language technologies for Celtic languages.

(2) **Joint Session with SLaTE 2023** (August 19)

SIGUL 2023 will have a Joint Session with the 9<sup>th</sup> Workshop on Speech and Language Technology in Education ([SLaTE 2023](#)). The goal is to accelerate the development of spoken language technology for under-resourced languages through education.

(3) **Social Outing and Dinner near Dublin** (August 20) [optional]

### Invited Speakers

- Subhashish Panigrahi, O Foundation and Law for All Initiative, Ashoka (Italy)  
*Reclaiming Our Voices: Imagining Community-Led Ai/ML Practices*
- Delyth Prys, School of Arts, Culture and Language, Bangor University | Canolfan Bedwyr (Wales, UK)  
*New Challenges, Old Problems: AI and Under-resourced Languages*

### Workshop Topics

Following the long-standing series of previous meetings, the SIGUL venue will provide a forum for the presentation of cutting-edge research in natural language processing and spoken language processing for under-resourced languages to both academic and industry researchers and 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 include but are not limited to:

- processing of any under-resourced language (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-resource speech technologies and self-supervised learning;
-

- cross-lingual and multilingual acoustic and lexical modeling;
- speech recognition and synthesis for under-resourced languages and dialects;
- machine translation and spoken dialogue systems;
- applications of spoken language technologies for under-resourced languages;
- special topics:
  - Celtic language technology;
  - spoken language technologies for under-resourced languages via education.

We also welcome various typologies of papers:

- research papers;
- position papers for reflective considerations of methodological, best practice, institutional issues (e.g., ethics, data ownership, speakers' community involvement, de-colonizing approaches);
- research posters for work-in-progress projects in the early stage of development or description of new resources;
- demo papers, and early-career/student papers, to be submitted as extended abstracts and presented as posters.

### Instructions for Submission

Prospective authors are invited to submit their contributions according to the following guidelines.

- Research and position papers: a maximum of 5 pages with the 5<sup>th</sup> page reserved exclusively for references.
- Demo papers and early-career/student papers: a maximum of 3 pages with the 3<sup>rd</sup> page reserved for references.

Both types of submissions must conform to the [Interspeech format](#) defined in the paper preparation guidelines as instructed in the [author's kit](#) on the Interspeech Webpage. Papers do not need to be anonymous. Authors must declare that their contributions are original and that they have not submitted their papers elsewhere for publication.

### Important Dates

- Paper submission deadline: 29 May 2023
-

- Notification of acceptance: 19 June 2023
- Camera-ready paper: 14 July 2023
- Workshop dates: 18-20 August 2023

All deadlines are 11:59 PM UTC-12:00 (“anywhere on Earth”).

## Outline of the Programme

SIGUL 2023 will continue the tradition of the previous SIGUL event 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.

## Full List of Organizers

### ***SIGUL Board***

Sakriani Sakti (JAIST, Japan)

Claudia Soria (CNR-ILC, Italy)

Maite Melero (BSC, Spain)

### ***SIGUL 2023 Organizing Committee***

Ailbhe Ní Chasaide (Phonetics and Speech Laboratory [abair.ie], TCD, Ireland)

Neasa Ní Chiaráin (Phonetics and Speech Laboratory [abair.ie], TCD, Ireland)

Emily Barnes (Phonetics and Speech Laboratory [abair.ie], TCD, Ireland)

Maite Melero (BSC, Spain)

Sakriani Sakti (JAIST, Japan)

Claudia Soria (CNR-ILC, Italy)

### ***SIGUL 2023 Programme Committee***

Gilles Adda (LIMSI/IMMI-CNRS, France)

Manex Agirrezabal (University of Copenhagen – Center for Sprogteknologi | Center for Language Technology, Denmark)

Shyam S. Agrawal (KIIT, India)

Begona Altuna (Euskal Herriko Unibertsitatea | University of the Basque Country, Spain)

Emily Barnes (Phonetics and Speech Laboratory and abair.ie, Trinity College Dublin, Ireland)

---

Steven Bird (Charles Darwin University, Australia)  
Matt Coler (University of Groningen, Campus Fryslân, The Netherlands)  
Pradip K. Das (IIT, India)  
Iria De Dios Flores ( Centro Singular de Investigación en Tecnoloxías Intelixentes, Spain)  
A. Seza Dođruöz (Universiteit Gent, België | Ghent University, Belgium)  
Stefano Ghazzali (Prifysgol Bangor | Bangor University, Bangor, Gwynedd)  
Jeff Good (University at Buffalo, USA)  
Kristiina Jokinen (AIRC [Artificial Intelligence Research Center], AIST Tokyo Waterfront, Japan)  
Laurent Kevers (Università di Corsica Pasquale Paoli, France)  
Teresa Lynn (Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates)  
Joseph Mariani (LIMSI-CNRS, France)  
Connor Mc Cabe (Phonetics and Speech Laboratory and abair.ie, Trinity College Dublin, Ireland)  
Maite Melero (Barcelona Supercomputing Center, Espanya | Spain)  
Andy Murphy (Phonetics and Speech Laboratory and abair.ie, Trinity College Dublin, Ireland)  
Ailbhe Ní Chasaide (Phonetics and Speech Laboratory and abair.ie, Trinity College Dublin, Ireland)  
Neasa Ní Chiaráin (Phonetics and Speech Laboratory and abair.ie, Trinity College Dublin, Ireland)  
Win Pa Pa (UCS Yangon, Myanmar)  
Delyth Prys (Prifysgol Bangor | Bangor University, Bangor, Gwynedd)  
Carlos Ramisch (Université Marseille, France)  
Sakriani Sakti (JAIST, Japan)  
Claudia Soria (CNR-ILC, Italia | Italy)  
Trond Trosterud (Norges Arktiske Universitet | The Arctic University of Norway)

## Acknowledgments

SIGUL is a joint Special Interest Group of the European Language Resources Association (ELRA) and of the International Speech Communication Association (ISCA). The SIGUL 2023 Workshop has been organized with the help of the local organizers of Interspeech 2023 and SLaTE 2023. This edition has been sponsored by Google and endorsed by Linguapax International.

To contact the organizers, please mail [sigul2023@ml.jaist.ac.jp](mailto:sigul2023@ml.jaist.ac.jp) (Subject: [SIGUL2023]).