# LREC 2020

## 12th Conference on Language Resources and Evaluation

### FIRST CALL FOR PAPERS

**Palais du Pharo**
**MARSEILLE, FRANCE**

**MAIN CONFERENCE: 13-14-15 MAY 2020**

**WORKSHOPS and TUTORIALS: 11-12-16 MAY 2020**

**Conference website: https://lrec2020.lrec-conf.org/**

The European Language Resources Association (ELRA) is glad to announce the 12th edition of LREC, organised with the support of national and international organisations among which AFCP, AILC, ATALA, CLARIN, ILCB, LDC, ....

**CONFERENCE AIMS**
LREC is the major event on Language Resources (LRs) and Evaluation for Human Language Technologies (HLT). LREC aims to provide an overview of the state-of-the-art, explore new R&D directions and emerging trends, exchange information regarding LRs and their applications, evaluation methodologies and tools, on-going and planned activities, industrial uses and needs, requirements coming from e-science and e-society, with respect to both scientific/technological issues as well as policy and organisational ones.

LREC provides a unique forum for researchers, industrials and funding agencies from across a wide spectrum of areas to discuss issues and opportunities, find new synergies and promote initiatives for international cooperation, in support of investigations in language sciences, progress in language technologies (LT) and development of corresponding products, services and applications, and standards.

**CONFERENCE TOPICS**
**Issues in the design, construction and use of LRs: text, speech, sign, gesture, image, in single or multimodal/multimedia data**
- Guidelines, standards, best practices and models for LRs interoperability
- Methodologies and tools for LRs construction and annotation
- Methodologies and tools for extraction and acquisition of knowledge
- Ontologies, terminology and knowledge representation
- LRs and Semantic Web (including Linked Data, Knowledge Graphs, etc.)
- LRs and Crowdsourcing
- Metadata for LRs and semantic/content mark-up

**Exploitation of LRs in systems and applications**
- Sign language, multimedia information and multimodal communication
- LRs in systems and applications such as: information extraction, information retrieval, audio-visual and multimedia search, speech dictation, meeting transcription, Computer-Aided Language Learning, training and education, mobile communication, machine translation, speech

translation, summarisation, semantic search, text mining, inferencing, reasoning, sentiment analysis/opinion mining, etc.
- Interfaces: (speech-based) dialogue systems, natural language and multimodal/multisensory interactions, voice-activated services, etc.
- Use of (multilingual) LRs in various fields of application like e-government, e-participation, e-culture, e-health, mobile applications, digital humanities, social sciences, etc.
- Industrial LRs requirements
- User needs, LT for accessibility

**LRs in the age of deep neural networks**
- Semi-supervised, weakly-supervised and unsupervised machine learning approaches
- Representation Learning for language
- Techniques for (semi-)automatically generating training data
- Cross-language NLP & Cross-domain NLP with reduction of human effort

**Issues in LT evaluation**
- LT evaluation methodologies, protocols and measures
- Validation and quality assurance of LRs
- Benchmarking of systems and products
- Usability evaluation of HLT-based user interfaces and dialogue systems
- User satisfaction evaluation

**General issues regarding LRs & Evaluation**
- International and national activities, projects and initiatives
- Priorities, perspectives, strategies in national and international policies for LRs
- Multilingual issues, language coverage and diversity, less-resourced languages
- Open, linked and shared data and tools, open and collaborative architectures
- Replicability and reproducibility issues
- Organisational, economical, ethical and legal issues

**LREC 2020 HOT TOPICS**

**Less Resourced and Endangered Languages**

Special attention will be devoted to less resourced and endangered languages: it is expected that LREC2020 makes room to activities carried out to support indigenous languages, building on the United Nations/UNESCO International Year of Indigenous Languages being celebrated in 2019.

**Language and the Brain**

Studying the neural basis of language helps in understanding both language processing and the brain mechanisms. LREC2020 will encourage all submissions addressing language and the brain. Among possible subtopics, submissions could focus on new datasets and resources (neuroimaging, controlled corpora, lexicons, etc.), methods aiming at new multimodal experimentations (e.g. EEG in virtual reality), language processing applications (e.g. brain decoding, brain-computer interfaces), etc.

**Machine/Deep Learning**

The availability of LRs is a key element of the development of high quality Human Language Technologies based on AI/Machine Learning approaches, and LREC is the best place to get access to this data, in many languages and for many domains. In addition to submissions addressing ML issues based on large quantities of data, those applied to languages for which only small, noisy or sparse data exist are also most welcomed.

**DESCRIBE AND SHARE YOUR LRs!**

In addition to describing your LRs in the LRE Map – now a normal step in the submission procedure of many conferences – LREC recognises the importance of sharing resources and making them available to the community.

When submitting a paper, you will be offered the possibility to share your LRs (data, tools, web-services,

etc.), uploading them in a special LREC repository set up by ELRA. Your LRs will be made available to all LREC participants before the conference, to be re-used, compared, analysed. This effort of sharing LRs, linked to the LRE Map for their description, contributes to creating a common repository where everyone can deposit and share data.

**PROGRAMME**
The Scientific Programme will include invited talks, oral presentations, poster and demo presentations, and panels, in addition to a keynote address by the winner of the Antonio Zampolli Prize.

We will also organise an **Industrial Track** and a **Reproducibility Track**: for these there will be separate Calls.

**SUBMISSIONS AND DATES**
Submission of **oral** and **poster** (or **poster+demo**) papers: **25 November 2019**
- LREC2020 asks for **full papers** from **4 pages to 8 pages (plus more pages for references if needed)**, which **must strictly follow the LREC stylesheet** which will be available on the conference website. Papers must be submitted through the LREC2020 submission platform (it uses START from Softconf) and will be peer-reviewed.

Submission of proposals for **workshops**, **tutorials** and **panels**: **24 October 2019**
- Proposals should be submitted via an online form on the LREC website and will be reviewed by the Programme Committee.

**PROCEEDINGS**
The **Proceedings will include both oral and poster papers**, in the same format.

**There is no difference in quality between oral and poster presentations. Only the appropriateness of the type of communication (more or less interactive) to the content of the paper will be considered.**

LREC 2010, LREC 2012 and LREC 2014 Proceedings are included in the Thomson Reuters Conference Proceedings Citation Index. The other editions are being processed.
LREC Proceedings are indexed in Scopus (Elsevier).

Substantially extended versions of papers selected by reviewers as the most appropriate will be considered for publication in a special issue of the Language Resources and Evaluation Journal published by Springer (a SCI-indexed journal).

**CONFERENCE PROGRAMME COMMITTEE**
Nicoletta Calzolari – CNR, Istituto di Linguistica Computazionale "Antonio Zampolli", Pisa - Italy
(Conference chair)
Frédéric Béchet – LIS-CNRS, Aix Marseille University, Marseille - France
Philippe Blache – CNRS & Aix-Marseille University, Marseille - France
Christopher Cieri – Linguistic Data Consortium, Philadelphia - USA
Khalid Choukri  – ELRA, Paris - France
Thierry Declerck – DFKI GmbH, Saarbrücken - Germany
Hitoshi Isahara – Toyohashi University of Technology, Toyohashi - Japan
Bente Maegaard  – Centre for Language Technology, University of Copenhagen, Copenhagen - Denmark
Joseph Mariani  – LIMSI-CNRS, Orsay - France
Asuncion Moreno  – Universitat Politècnica de Catalunya, Barcelona - Spain
Jan Odijk  – UIL-OTS, Utrecht - The Netherlands

Stelios Piperidis – Athena Research Center/ILSP, Athens - Greece

**CONFERENCE EDITORIAL COMMITTEE**

Sara Goggi –  CNR, Istituto di Linguistica Computazionale "Antonio Zampolli", Pisa - Italy
Hélène Mazo –  ELDA/ELRA, Paris - France