

GWLL – GLOBALEX Workshop on Linked Lexicography @ LREC 2022 – Call for Papers

Background. The 4th GLOBALEX Workshop at LREC will elaborate the theme of linking data across lexicographic resources and with other lexical resources, in the aim of enhancing language data methodologies and applications.

Linking lexicographic data sets to each other and to other language resources, and in particular the interoperability of lexicography with Linked Data (LD) technologies, have gained substantial attention in recent years, becoming a subject of projects for research and collaboration between academia and industry, including support of the public sector. Most notably, the W3C community group on Ontology-Lexica ([OntoLex](#)) was formed with the mission to “develop models for the representation of lexica (and machine-readable dictionaries) relative to ontologies”. The ensuing [OntoLex-lemon](#) model [2] has served since 2016 as the leading option for conversion of lexicographic data into LD, and was updated with the [lexicog](#) module released in [2019](#).

Description. Besides general research papers, GWLL 2022 will include two tracks combined with shared tasks:

- (1) Linking Monolingual Lexicographic Resources – in conjunction with a shared task conducted by ELEXIS.
- (2) Linking Bilingual and Multilingual Lexicographic Resources – in conjunction with the 5th TIAD [shared task](#).

Task 1 will be evaluated on novel dictionary linking data developed by the [ELEXIS](#) project and cover linking for the following languages: Bulgarian, Danish, Dutch, English, Estonian, German, Hungarian, Irish, Italian, Slovene and Spanish.

Task 2 will host the 5th edition of the Translation Inference Across Dictionaries (TIAD) shared task – whose previous editions were co-located at Language, Data and Knowledge conferences [13], [14], [15], and at LREC 2020 [16]) – with the aim to explore methods and techniques for automatically generating new bilingual (and multilingual) dictionaries from existing ones, including validation of results and comparison of the processes used. The participating systems are invited to generate new translations automatically among English, French and Portuguese (and other languages are expected to be added during the preparation phase), based on known translations contained in the [Apertium](#) RDF graph.

Background. This full-day event constitutes the fourth iteration in the successful GLOBALEX @ LREC workshop series – following up on the first Workshop on Lexicographic Resources and Human Language Technology at LREC [2016](#) and the second Workshop on Lexicography & Wordnets at LREC [2018](#) (as well as the third Workshop on Linked Lexicography that was due as part of LREC 2020). GLOBALEX Workshop

@ LREC 2022 is supported by the Global Alliance for Lexicography ([Globalex](#)), ELEXIS, [NexusLinguarum](#) – European network for Web-centred linguistic data science, COST Action CA18206, and TIAD.

Topics. We invite proposals for papers and demonstrations on the main topic of linking lexicographic resources, as well as on standards for the development, representation, evaluation and dissemination of lexicographic resources, and the interoperability of lexicography with other disciplines and external resources and domains including computer-assisted language learning and translation, neural machine translation, terminology, multilinguality, natural language processing and understanding, automatic language generation, linguistic linked (open) data and semantic web technologies, knowledge systems, deep learning and (explainable) artificial/augmented intelligence.

- Linking monolingual dictionaries and lexicographic resources
- Linking bilingual dictionaries and lexicographic resources
- Linking multilingual dictionaries and lexicographic resources
- Linking lexicographic data with other lexical data resources
- Applications and developments of the OntoLex-*lemon* model and Lexicog module
- RDF serializations of lexicographic data
- Non-RDF data formats for linked lexicographic resources
- RDF and XML standards for linked lexicography
- Converting lexicographic data for linking purposes
- Linked Data-native lexicographic resources
- Automated generation of lexicographic resources based on Linked Data technologies
- Lexicography, terminology and Linguistic Linked (Open) Data
- Linked lexicography and the Semantic Web
- Linked lexicography and the Multilingual Digital Single Market
- Linked lexicography and Knowledge Systems
- Linked lexicography and Artificial/Augmented Intelligence
- Linked lexicography, deep learning and neural networks

Submission and Timetable

We invite you to submit a 1,000-word abstract. Accepted proposals will be invited to submit a full (short or long) paper to be included in the LREC proceedings. Templates and author instructions are available [here](#). Please note these important dates:

- **April 8, 2022** Abstract submission deadline
- **April 29, 2022** Notification
- **May 27, 2022** Full paper submission deadline

Programme Committee

Workshop Organizers:

- Ilan Kernerman, K Dictionaries – Lexicala
- Simon Krek, Globalex, Jožef Stefan Institute

Track 1 Organizers:

- John McCrae, National University of Ireland Galway
- Sina Ahmadi, National University of Ireland Galway

Track 2 Organizers:

- Jorge Gracia, University of Zaragoza
- Besim Kabashi, Friedrich-Alexander University of Erlangen-Nuremberg

Review Committee (tentative)

- Andreas Blumauer, Semantic Web Company, Vienna
- Michal Boleslav Méchura, Masaryk University
- Nicoletta Calzolari, Computational Linguistics Institute, Pisa
- Christian Chiarcos, Goethe University
- Philipp Cimiano, University of Bielefeld
- Thierry Declerck, University of Saarbrücken
- Gerard de Melo, Rutgers University
- Tomaz Erjavec, Jožef Stefan Institute
- Christiane Fellbaum, Princeton University
- Edward Finegan, University of Southern California
- Radovan Garabik, L. Štúr Institute of Linguistics, Bratislava
- Alexander Geyken, Berlin-Brandenburg Academy of Sciences and Humanities
- Asunción Gómez Pérez, Madrid Polytechnic University
- Sebastian Hellmann, Leipzig University
- Isahara Hitoshi, Toyohashi University of Technology
- Ales Horak, Masaryk University
- Chu-Ren Huang, Hong Kong Polytechnic University
- Miloš Jakubiček, Lexical Computing
- Jelena Kallas, Institute of the Estonian Language, Tallinn
- Kyoko Kanzaki, Toyohashi University of Technology
- Diptesh Kanojia, IIT Bombay
- Iztok Kosem, University of Ljubljana
- Lothar Lemnitzer, Berlin-Brandenburg Academy of Sciences and Humanities
- Nikola Ljubešić, University of Zagreb
- Stella Markantonatou, Institute for Language and Speech Processing (ATHENA)
- Verginica Mititelu, Romanian Academy Research Institute for Artificial Intelligence
- Roberto Navigli, Sapienza University
- Adam Pease, Articulate Software
- Bolette Pedersen, University of Copenhagen
- Maciej Piasecki, Wrocław University of Technology
- Alexandre Rademaker, IBM Research – Brazil
- Egon Stemle, European Academy of Bozen/Bolzano
- Lars Trap Jensen, Society for Danish Language and Literature

- Tamás Váradi, Hungarian Academy of Sciences
- Kadri Vider, University of Tartu
- Piek Vossen, Vrije Universiteit Amsterdam

Identify, Describe, and Share your LRs!

Describing your LRs in the LRE Map is now a normal practice in the submission procedure of LREC (introduced in 2010 and adopted by other conferences). To continue the efforts initiated at LREC 2014 about “Sharing LRs” (data, tools, web-services, etc.), authors will have the possibility, when submitting a paper, to upload LRs in a special LREC repository. This effort of sharing LRs, linked to the LRE Map for their description, may become a new “regular” feature for conferences in our field, thus contributing to creating a common repository where everyone can deposit and share data.

As scientific work requires accurate citations of referenced work so as to allow the community to understand the whole context and also replicate the experiments conducted by other researchers, LREC 2022 endorses the need to uniquely identify LRs through the use of the International Standard Language Resource Number (ISLRN, www.islrn.org), a Persistent Unique Identifier to be assigned to each Language Resource. The assignment of ISLRNs to LRs cited in LREC papers will be offered at submission time.