LANGUAGE, DATA and KNOWLEDGE 2025

Proceedings of the 5th Conference on **Language, Data and Knowledge: Workshops**

PROCEEDINGS OF THE 5TH CONFERENCE ON LANGUAGE, DATA AND KNOWLEDGE: WORKSHOPS

EDITORS:

Katerina Gkirtzou, ILSP "Athena" Research Center, Greece
Slavko Žitnik, University of Ljubljana, Slovenia
Jorge Gracia, University of Zaragoza, Spain
Dagmar Gromann, University of Vienna, Austria
Maria Pia di Buono, University of Naples "L'Orientale", Italy
Johanna Monti, University of Naples "L'Orientale", Italy
Maxim Ionov, University of Zaragoza, Spain

UniorPress

Table of Contents

Fifth OntoLex Workshop

Fifth OntoLex Workshop

Fifth OntoLex Workshop

The Fifth OntoLex workshop, which takes place in Naples, Italy on September 9th, 2025, is the latest in a series of events dedicated to the OntoLex-Lemon vocabulary and which has previously been held in conjunction with the "Language, Data and Knowledge" conference series in Leipzig (2019), Zaragoza (2021), and Vienna (2023) and as a stand-alone event in Leiden (2018).

Having established itself as a pivotal standard for representing lexical data as linked data, the use of the OntoLex-Lemon model has recently expanded into diverse domains, including computational lexicography, terminology management, and AI-driven language technologies. However, as new linguistic challenges and new technological advancements emerge, it is imperative that the OntoLex model continue evolving to meet these needs. With this in mind, the aim of the OntoLex workshop is to provide a forum for sharing ideas and proposals in order to meet these new challenges and to ensure that OntoLex keeps pace with the latest developments. Overview: In the morning session, the workshop features a series of invited talks by experts in the area, this is followed by six contributed papers which have been peer-reviewed by the program committee (see below); these are as follows. The paper Inferring Adjective Hypernyms with Language Models"(Augello and McCrae) explores how masked language models can detect semantic relations like hypernymy between adjectives, leveraging Open English WordNet. Philosophising Lexical Meaning as an OntoLex-Lemon Meta-Ontology"(Zamborlini, Zhu, van Erp and Betti) introduces a conceptual framework to extend OntoLex-Lemon on a philosophically informed basis. Bringing IATE into the Semantic Web Family"(Diez-Ibarbia, Martín-Chozas and Montiel-Ponsoda) reports on ongoing efforts to convert the EU's Interactive Terminology for Europe (IATE) into an RDF-compliant resource, contributing structured terminology to the Linked Open Data cloud. OntoLex-Lemon in Wikidata and other Wikibase instances"(Lindemann) discusses the mapping of lexical data to Wikidata entities using OntoLex-Lemon, describing methods for aligning linguistic annotations and glosses with multilingual data stores, including practical challenges in using Wikibase software. The article A Lightweight String Based Method of Encoding Etymological Information in RDF"(Khan, Ionov, Marongiu and Salgado) presents a strategy for representing etymological content as RDF string literals using a well-defined regular grammar. This method supports a shallow description of lexical histories that can be queried using the SPARQL regular expression function. Finally Ontologies for historical languages: using the LiLa and OntoLex-Lemon framework to build a Lemma Bank for Old Irish"(Fransen) focuses on historical language data, showing the use of OntoLex in developing a lemma bank for Old Irish, similar to the one created in the Linking Latin project.

The afternoon session of the workshop, which is also open to remote participation by the members of the W3C Ontology-Lexicon group, provides an overview of the latest advancements in the OntoLex model. It features an introduction to preparatory work on a forthcoming 2.0 version of OntoLex, status updates on the development of the FrAC and Morph modules, software demos and proposals for new modules, e.g. for etymology and conceptual modelling.

<div align="right">

Fahad Khan, John McCrae, Matteo Pellegrini and Philipp Cimiano
The OntoLex 2025 Organising Committee

</div>

Organizing Committee

Anas Fahad Khan, Istituto di Linguistica Computazionale A. Zampolli, Consiglio Nazionale delle Ricerche, Pisa, Italy
John Philip McCrae, Data Science Institute, National University of Ireland Galway, Galway, Ireland
Matteo Pellegrini, Università Cattolica del Sacro Cuore, Milan, Italy
Philipp Cimiano, Bielefeld University, Bielefeld, Germany

Program Committee

Maxim Ionov, University of Zaragoza
Sina Ahmadi, University of Zurich
Patricia Martín-Chozas, Universidad Politécnica de Madrid
Katerina Gkirtzou, ILSP "Athena" Research Center, Greece
Gilles Sérasset, LIG - Université Grenoble Alpes
Francis Bond, Palacký University
Christian Chiarcos, University of Augsburg
Ana Salgado, NOVA Lisbon
Elena Montiel-Ponsoda, Universidad Politécnica de Madrid
Adrian Doyle, University of Galway
Paul Buitelaar, University of Galway
Sara Carvalho, University of Aveiro
Jorge Gracia, University of Zaragoza
Manuel Fiorelli, University of Rome, Tor Vergata