

Call for Papers

# 6th Workshop on Linked Data in Linguistics (LDL-2018): Towards Linguistic Data Science

Saturday, 12 May 2018, Miyazaki, Japan, <http://ldl2018.linguistic-lod.org/>, held in conjunction with the 11th Language Resources and Evaluation Conference (LREC-2018)

The Linked Data in Linguistics (LDL) workshop series has become the major forum for presenting, discussing and the application of **Semantic Web standards and the Linked Open Data paradigm to language resources** and will take place again in 2018 in co-location with the 11th Language Resources and Evaluation Conference in Miyazaki, Japan. The LDL workshop series has a general focus on LOD-based language resources, vocabularies, infrastructures and technologies as means for **managing, improving and using language resources on the Web**, addressing communities as diverse as applied linguistics, lexicography, digital humanities, natural language processing and information technology. LDL provides the primary forum for these communities **to present and to discuss use cases, experiences, best practices, recommendations and technologies** of Linguistic Linked Open Data in an interdisciplinary setting, addressing topics such as

- ❑ **linking language resources:** Linked Data facilitates to the interoperability, re-usability and discoverability of language resources as well as information integration across them.
- ❑ **natural language for structured knowledge:** Ontologies and knowledge bases in general are often created with one particular language in mind. Their lexicalization for other natural languages is a prime directive for emerging digital infrastructures in a globalized world.
- ❑ **structured knowledge for natural language:** Natural language is in constant development, thus inherently imprecise. Grounding natural language text in structured knowledge (word sense disambiguation, entity linking, semantic parsing) enables machine readability and processability by downstream tasks.

Moreover, as technology and resources increasingly converge towards a LOD-based ecosystem for the systematic cross-lingual discovery, exploitation, extension, and curation of language resources, we particularly encourage contributions focusing on the development of **research methodologies and applications** building on the Linguistic Linked Open Data Cloud and the existing technology and resource stack: The notion of **Linguistic Data Science** brings together Linguistic Linked Open Data with research questions in linguistics, methods in Natural Language Processing and applications in Digital Humanities. The workshop will also provide the opportunity for the **critical reflection** of Linked Open Data techniques and their application in linguistics, natural language processing and related fields, i.e., to express experiences, challenges, crucial issues, missing aspects regarding the reuse, creation and exploitation of Linguistic Linked Open Data.

## Topics of Interest

We invite presentations of algorithms, methodologies, experiments, use cases, descriptions of research projects and position papers regarding the creation, publication or application of language resources and their linking, as well as descriptions of such data and its uses in research (linguistics, lexicology, Digital Humanities) and technology (natural language processing, lexicography, localization). This includes the following aspects:

- Building and managing linked language resources
  - Vocabularies and best practices for language resources and their linking.
  - Application of LOD in language resource infrastructures.
  - Metadata linking and curation for language resources on the Web.
- LLOD technology and methodology
  - Methodologies to develop linked language resources on the Web.
  - Using natural language processing to enhance Linked Open Data.
  - Leveraging Linguistic Linked Open Data and Machine Learning.
  - Methods for linking dictionaries across languages.
- LLOD applications
  - Using Linked Open Data to facilitate natural language processing.
  - Using Linked Open Data for quantitative and qualitative linguistic research.
  - Linguistic Linked Open Data for Digital Humanities.
  - Linking language resources and external community resources.
  - Using Linked Open Data to connect lexicographic resources.
- Critical Reflection:
  - Challenges of scalability, multilinguality and interoperability in the Web.
  - Legal, social and scientific aspects of Linguistic Linked Open Data.
  - LLOD revisited: Case studies, use cases, and lessons learnt.

We invite both long (8 pages and 2 pages of references, formatted according to the LREC guidelines) and short papers (4 pages and 2 pages of references) representing original research, innovative approaches and resource descriptions. Short papers may also represent project descriptions. These do not have to be implemented but discuss to what extent and for which purposes Linguistic Linked Open Data is reused or created. Projects that are still in their early stages and seek advice from the broader Linguistic Linked Data community are welcome, especially if they include underrepresented fields of study.

## Important Dates

12 Jan 2018: submission  
 15 Feb 2018: notification  
 02 Mar 2018: camera-ready  
 12 May 2018: workshop in  
 Miyazaki (Japan)

## Organizing Committee

Christian Chiarcos (Goethe University Frankfurt, Germany)  
 John P. McCrae (National University of Ireland Galway)  
 Thierry Declerck (University of Saarland, Germany)  
 Jorge Gracia (University of Zaragoza, Spain)  
 Bettina Klimek (University of Leipzig, Germany)

Please get in contact via [ldl2018@linguistic-lod.org](mailto:ldl2018@linguistic-lod.org)

## Identify, Describe and Share your LRs!

Describing your LRs in the LRE Map is now a normal practice in the submission procedure of LREC. 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 2018 endorses the need to uniquely identify LRs through the use of the International Standard Language Resource Number (ISLRN, [www.islrn.org](http://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.

## Programme Committee

- Eneko Agirre (University of the Basque Country, Spain)
- Guadalupe Aguado-de-Cea (Universidad Politécnica de Madrid, Spain)
- Núria Bel (Universitat Pompeu Fabra, Spain)
- Claire Bonial (University of Colorado at Boulder, USA)
- Paul Buitelaar (Insight Center for Data Analytics, National University of Ireland Galway, Ireland)
- Nicoletta Calzolari (ILC-CNR, Italy)
- Steve Cassidy (Macquarie University, Australia)
- Damir Cavar (Indiana University, USA)
- Philipp Cimiano (University of Bielefeld, Germany)
- Gerard de Melo (Rutgers University, USA)
- Francesca Frontini (Université Paul Valéry, Montpellier, France)
- Jeff Good (University at Buffalo, USA)
- Dagmar Gromann (IIIA-CSIC, SPAIN)
- Yoshihiko Hayashi (Osaka University, Japan)
- Fahad Khan (ILC-CNR, Italy)
- Dave Lewis (ADAPT, Ireland)
- Vanessa Lopez (IBM Research, Ireland)
- Elena Montiel-Ponsoda (Universidad Politécnica de Madrid, Spain)
- Steve Moran (Universität Zürich, Switzerland)
- Roberto Navigli ("La Sapienza" Università di Roma, Italy)
- Sebastian Nordhoff (Language Science Press, Berlin, Germany)
- Petya Osenova (IICT-BAS, Bulgaria)
- Antonio Pareja-Lora (Universidad Complutense Madrid, Spain)
- Francesca Quattri (Jiangsu University, China)
- Mariano Rico (Universidad Politécnica de Madrid, Spain)
- Laurent Romary (INRIA, France)
- Felix Sasaki (Berlin, Germany)
- Andrea Schalley (Karlstad University, Sweden)
- Gilles Sérasset (University Grenoble Alpes, France)
- Milena Slavcheva (JRC-Brussels, Belgium)
- Armando Stellato (University of Rome, Tor Vergata, Italy)
- Marieke van Erp (KNAW Humanities Cluster, the Netherlands)
- Cristina Vertan (University of Hamburg, Germany)
- Piek Vossen (Vrije Universiteit Amsterdam, The Netherlands)