8th Workshop on Linked Data in Linguistics (LDL-2022): Revisiting a Decade of Linguistic Linked Open Data

Workshop Description

Since its inception, the workshop series on Linked Data in Linguistics (LDL) established itself as the main venue for discussing how Linked Open Data (LOD) and semantic web technologies can be used for processing, analysing, publishing, and managing linguistic data. This includes the fields of natural language processing (NLP), language resources (LRs), lexicography and digital humanities (DH), and has been leading to the development of **linguistic data science as** a new area of study. The LDL workshop series has contributed greatly to the development of the **Linguistic Linked Open Data (LLOD)** cloud and the development of best practices for publishing and accessing language resources and providing language technology services on the web. Most notably, this includes community standards such as the NLP Interchange Format (NIF), the OntoLex-Lemon model of the W3C Community Group *Ontology-Lexica,* and numerous domain-specific adaptations and extensions that these models have had an influence on.

In addition, there are an increasing number of national, European, and international research projects that build on LLOD technology. These will contribute to its further development and will help ensure the success of this workshop and a high attendance rate. The 10th anniversary edition of the LDL can count on the support of the COST action "NexusLinguarum: European Network for Web-centered Linguistic Data Science", as well as two Horizon 2020 projects. Firstly, the *Prêt-à-LLOD* project, which is making linguistic linked open data ready-to-use, and, secondly, the *ELEXIS* project on building a lexicographic infrastructure.

Motivation and Topics of Interest

The LDL workshop series invites presentations of algorithms, methodologies, experiments, tools, use cases, descriptions of ongoing or planned research projects as well as position papers regarding the creation, support, publication or application of linguistic data collections and their linking with other resources. Descriptions of such data, and in particular, its uses in research (linguistics, lexicology, DH) and technology (natural language processing, lexicography, localization) are also welcome.

Additionally, the 10th anniversary edition of LDL will specifically welcome contributions and position papers that reflect on the impact, prospects, requirements or challenges encountered with the application of LLOD technology to research problems and industrial applications in the past decade. In particular, this includes experiences in the development, usage or limitations of existing and emerging standards as well as the application and adaptation of vocabularies for novel applications and domains, as well as contributions that look towards the future of LLOD technologies and applications for academic and industrial research.

We invite contributions focusing on, but not restricted to the following aspects of creating and managing Linked LRs:

1. Infrastructure for building & managing linked LRs
    1. Vocabularies, standards and best practices
    2. Tools for creating and curating LL(O)D resources
    3. Approaches for creating linked LRs
2. LLOD technology and methodology
    1. Methodologies for linked language resource development
    2. Leveraging knowledge graphs and machine learning

3. Linking resources across languages and domains
3. LLOD applications in different communities of practice, e.g., using
   1. LLOD to facilitate NLP and language technology applications,
   2. LLOD for quantitative and qualitative linguistic research,
   3. LLOD for lexicography, philologies and other branches of the humanities
4. Future directions and critical reflection
   1. Bringing LLOD to the end user: Building and using the technical ecosystem
   2. Ethical, legal, social and scientific aspects of Linguistic Linked Open Data
   3. Achieving FAIRness for language resources and NLP tools

Projects in their early stages that seek advice from the broader LLOD community are very welcome, especially if addressing fields still underrepresented in the LLOD cloud, such as the study of *low resource languages* or matters of *multimodality*.

Technical Requirements

A projector and poster boards will be required for presentation of the papers and posters. Technical support for hybrid event (playing recorded presentations, 2-3 microphones, Zoom connection, etc.).

Organizers

- John P. McCrae (National University of Ireland Galway)
- Thierry Declerck (DFKI GmbH)
- Elena Montiel (Universidad Politécnica de Madrid)
- Christian Chiarcos (Goethe University Frankfurt)
- Maxim Ionov (Goethe University Frankfurt)

Email address for contact

ldl2022@linguistic-lod.org

Programme Committee

New PC Members, all confirmed

- Sina Ahmadi (NUI Galway, Ireland)
- Sara Carvalho (University of Aveiro, Portugal)
- Milan Dojchinovski (Czech Technical University in Prague, Czech Republic)
- Agata Filipowska (Poznan University of Economics, Poland)
- Eero Hyvönen (Aalto University, Finland)
- Mietta Lennes (University of Helsinki, Finland)
- Chaya Liebeskind (Jerusalem College of Technology, Israel)
- Verginica Mititelu (Research Institute for Artificial Intelligence of the Romanian Academy, Romania)
- Ana Ostroški Anić (Institute of Croatian Language and Linguistics, Croatia)
- Ranka Stankovic (University of Belgrade, Serbia)
- Andrius Utka (Vytautas Magnus University, Lithuania)

PC Members of the past conference (to be confirmed)

- Eneko Agirre (University of the Basque Country, Spain)
- Núria Bel (Universitat Pompeu Fabra, Spain)
- Paul Buitelaar (Insight, Ireland)
- Nicoletta Calzolari (ILC-CNR, Italy)
- Steve Cassidy (Macquarie University, Australia)
- Philipp Cimiano (University of Bielefeld, Germany)

- Marieke van Erp (KNAW Humanities Cluster, the Netherlands)
- Francesca Frontini (Université Paul-Valéry, Montpellier, France)
- Jeff Good (University at Buffalo, USA)
- Dagmar Gromann (Vienna University, Austria)
- Yoshihiko Hayashi (Osaka University, Waseda University, Tokyo, Japan)
- Fahad Khan (ILC-CNR, Italy)
- Bettina Klimek (University of Leipzig, Germany)
- Dave Lewis (ADAPT, Ireland)
- Vanessa Lopez (IBM Research, Ireland)
- Gerard de Melo (HPI/Universität Potsdam, Germany)
- Steve Moran (University of Neuchâtel, 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 de Alcalá de Henares)
- Laurent Romary (INRIA, France)
- Felix Sasaki (SAP, Germany)
- Andrea Schalley (Karlstad University, Sweden)
- Gilles Sérasset (University Grenoble Alpes, France)
- Milena Slavcheva (IICT-BAS, Bulgaria)
- Armando Stellato (University of Rome, Tor Vergata, Italy)
- Piek Vossen (Vrije Universiteit Amsterdam, The Netherlands)