

# OntoLex-2017: 1st Workshop on the OntoLex Model

**John P. McCrae♠, Elena Montiel Ponsoda♥, Philipp Cimiano◇,**

**Paul Buitelaar♠, Jorge Gracia♥ and Thierry Declerck♣**

♠Insight Centre for Data Analytics, National University of Ireland Galway

♥Universidad Politécnica de Madrid

◇CIT-EC, Bielefeld University

♣Saarland University

The Lemon-OntoLex model<sup>1</sup> has been developed for the representation of lexicons relative to ontologies on the Semantic Web, by the W3C OntoLex Community Group<sup>2</sup>, and has become the primary method for representing linked lexical data on the Web of Data. This model provides a new opportunity to bridge the gap between linguistics and NLP practitioners and the data science community by making large amounts of high quality lexical information available and linked to the knowledge that is explicitly represented on the Semantic Web, in graphs such as DBpedia. The Lemon-OntoLex model is finding applications in many areas. For example the model is being used to link wordnets, as part of the Global WordNet Association Collaborative Interlingual Index and furthermore in the representation of existing lexical resources including BabelNet, DBnary and commercial dictionaries.

As such, this represents a crucial point for the various users of the model to come together and discuss the current applications of the model as well as future developments of the model. For this reason, we established a new workshop series on the OntoLex model and its applications, that was co-located with the First Conference on Language, Data and Knowledge (LDK 2017) in Galway, Ireland.

We accepted nine papers to this workshop all of which received very positive reviews and these were grouped into three sessions on “Tools and Experiences”, “Applications” and “Extensions”. The first session, included a paper by Declerck outlining the challenges in using the Lemon-OntoLex model for standard dictionaries, and a second paper by Chiarcos et al., discussed the particular case of applying the model to Turkic languages, showing the value of the model for linguists working with minority languages. Then Fiorelli et al. presented their work on building an API to work with the LIME module, a part of the OntoLex specification.

The second session showed something of the applications of the model, firstly with Francesconi and Hajlaoui describing the application of the model at the EU publication office’s knowledge management infrastructure. Khan et al. described the challenges of using

---

<sup>1</sup> <http://www.w3.org/2016/05/ontolex/>

<sup>2</sup> <https://www.w3.org/community/ontolex/>

the model for historical languages, in particular by converting the Liddell-Scott Lexicon of Ancient Greek into the OntoLex model. Finally, Balysheva et al. showed the application of the model and the use of the LexInfo categories to Russian.

The final session, looked at novel extensions of the model. Bosque-Gil et al. proposed a new lexicographic model to enable the model to better represent dictionaries such as those seen in earlier talks. Stolk expanded on this, focussing on the onomasiological structuring of concepts in OntoLex and further suggesting how concepts can be organized in the model. Finally, Klimek proposed developing a module for morphological representation based on her existing MMoOn ontology.

The paper presentations were followed by a lively discussion that has brought fresh impetus to the work on the model and will restart the development of the model around four new modules on Lexicography, Morphology, Etymology and Lexical Categories. We expect the next instance of the workshop to further expand on these goals as well as continue to support applications and tools for the model.

## Organization

### *Organizing Committee*

John P. McCrae - Insight Centre for Data Analytics, National University of Ireland, Galway  
Elena Montiel Ponsoda - Universidad Politécnica de Madrid  
Philipp Cimiano - CIT-EC, Bielefeld University  
Paul Buitelaar - Insight Centre for Data Analytics, National University of Ireland, Galway  
Jorge Gracia - Universidad Politécnica de Madrid  
Thierry Declerck - Saarland University

### *Program Committee*

Guadalupe Aguado-de-Cea - Universidad Politécnica de Madrid  
Paul Buitelaar - Insight Centre for Data Analytics, National University of Ireland, Galway  
Christian Chiarcos - Goethe-University Frankfurt  
Philipp Cimiano - CIT-EC, Bielefeld University  
Thierry Declerck - Saarland University  
Francesca Frontini - Université Paul Valéry - Montpellier 3  
Jorge Gracia - Universidad Politécnica de Madrid  
Bettina Klimek - InfAI, University of Leipzig  
Fahad Khan - ILC-CNR  
John P. McCrae - Insight Centre for Data Analytics, National University of Ireland, Galway  
Elena Montiel Ponsoda - Universidad Politécnica de Madrid  
Laurette Pretorius - Unisa  
Armando Stellato - University of Rome, Tor Vergata  
Felix Sasaki - W3C / DFKI  
Piek Vossen - Vrije Universiteit Amsterdam