Standardizing syntactically rich NLP lexicons for the Semantic Web: use cases of the lemon lexicon model

Judith Eckle-Kohler (Technische Universität Darmstadt)

Based on joint work with John Philip McCrae and Christian Chiarcos.

Standardizing NLP lexicons - representing them in a common representation format - makes them interoperable and plays an important role in the context of lexicon integration and reuse. In my talk I will present a mapping between two standard formats for encoding NLP lexicons: the UBY-LMF lexicon model, an instantiation of the abstract standard ISO Lexical Markup Framework, and the lemon lexicon model, which has been designed for ontology lexicons on the Semantic Web.

Both UBY-LMF and lemon have been used to standardize a broad range of lexical resources, for example wordnets and Wiktionary, but also syntactically rich lexicons such as the PAROLE lexicons (lemon) or a subset of IMSLex (UBY-LMF).

I will put a special focus on subcategorization frames and lexical syntax and discuss how lemon and UBY-LMF represent these information types.