1st LIDER Roadmapping Workshop

21st of March, 2014

»Report of Ontolex W3C Community Group Activities«
The lexical gap

http://dbpedia.org/ontology/spouse

How is this verbalized in different languages?
The lexical gap

- English: married to, husband of, wife of, ...
- German: verheiratet mit, Frau von, Mann von, ...
- Spanish: casado con, marido de, mujer de, ...
- ...

Philipp Cimiano
Aim: provide rich linguistic grounding for domain ontologies

Principles:

- Model of lexicon–ontology interface
- m:n relation between lexicon entries / lexica and ontologies
- Divide–and–Conquer / Open World
- Move away from labels! –> Lexical Entities / Linguistic Objects
- Principle of semantics by reference
- Theory–agnostic, top level only!
- Semantic Web / Linked Data Principles
History

- 2006: LingInfo: Buitelaar, Sintek et al. (DFKI)
- 2006: LIR: Asunción Gómez Pérez, Lupe Aguado de Cea, Elena Montiel Ponsoda (UPM, NeON Project)
- 2007: LexOnto: Cimiano, Buitelaar
- 2009: LexInfo: Buitelaar, Cimiano, McCrae (SmartWeb Project)
- 2010: lemon: lexicon model for ontologies (Monnet project)
- since 2012: ontolex W3C Community Group (specification expected in 2014)
Exploitation

Semantic NLP services:

- Entity Linking (Semantic Normalization)
- Question Answering over Linked / RDF Data
- Natural Language Generation from RDF Data
- ...
Interoperability

- Syntactic
- Semantic

Syntactic interoperability follows from semantic one
5 Modules

- Ontology-lexicon interface (ontolex)
- Syntax and semantics (synsem)
- Decomposition (decomp)
- Variation and translation (var)
- Linguistic Metadata (lime)
State of Play

- Chairs: Paul Buitelaar, Philipp Cimiano
- Participants: 78
- After major debates in 2013, consensus established in 2014
- First version of model to be released in summer (5 modules)
Model

- APIs
- validation services
- reference applications that exploit the model
Power is in the linking

- ontolex model as building block for a network of linked lexical resources, following Linked Data principles
- not one effort, but distributed and linked
- model for Global WordNet ???
Looking for

- real-world use cases
- NLP services exploiting the lexicon ontology model
- users and adopters
- people instantiating / populating it