

publishing and consuming lexicographical resources in the linked (open) data cloud



THIERRY DECLERCK 1
THIERRY.DECLERCK@DFKI.DE

EVELINE WANDL-VOGT 2
EVELINE.WANDL-VOGT@OEAW.AC.AT



1 deutsches forschungszentrum für künstliche intelligenz (DFKI)
+ saarland university, department for computational linguistics

2 austrian academy of sciences (ÖAW), austrian center for digital humanities (ACDH)
+ austrian standards institute (ASI)

motivation

lexicography is at a crossroads.

the digital turn, followed by the semantic turn, did not just mean a change of format for lexicographers, linguists and terminologists, but implies a complete turn-over of the workflow.

to be a lexicographer, linguist or terminologist maybe never was as ambiguous as now.

in this tutorial we would like to open up views to a specific up-to-date topic of new possibilities to exploit potentials for lexicographical resources and the people connected with that:

linked open data (LOD).

linked data is a term used to describe a recommended best practice for exposing, sharing and connecting pieces of data, information and knowledge on the semantic web [wikipedia].

by interlinking multilingual and open (language) resources, we foresee a linguistic linked open data (LLOD) cloud, a new linguistic ecosystem, that will allow the open exploitation of such data.

tutorial description:

the tutorial will give insights in the essentials of linked open data (LOD) and how those can be related to lexicographical (and other language) data.

the tutorial will contain a **linked open data start guide** for both

- a) publishing and
- b) consuming lexicographical data in the LOD.

we describe the potentials of linked open data for lexicography by means of concrete examples, dealing with porting of (dialectal) dictionaries into the LOD framework.

the tutorial will cover the following topics:

1. foundations

we will provide the necessary foundations in linked data (LD) and knowledge representation on the web (ontologies, RDF, SKOS, etc) in order to undertake practical work in the rest of the tutorial.



supported by



in connection with



2. modelling lexicographic resources in the linked data with RDF, SKOS and lemon

the second working unit aims at giving an overview of formal representation languages that can be used for modeling language data in the broader context of the semantic web.

we will give concrete examples on how to publish lexicographic resources in the LOD. we will illustrate how lexicographic data can be linked to other language data but also to domain ontologies representing specific knowledge in different fields.

3. methodology for publishing lexicographic data in the LOD

we focus on the process of generating and publishing LOD compliant lexicographic data from existing electronic dictionary resources.

the goal is to present the main steps of the LD generation process and some tools that can be used for LD generation and linking.

4. hands-on

participants will be given ample space to test the encoding of lexicographical data in well known, freely available ontology editors.

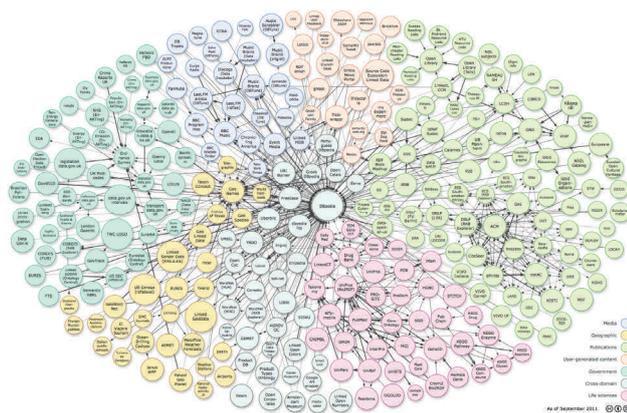
please bring your laptop and install beforehand

—: protégé (<http://protege.stanford.edu/>)

—: topbraid composer (<http://topbraid-composer-free-edition.software.informer.com/>).

facts

- tuesday,
july 15th 2014
- 09:00-13:00
- language: english
- EURAC research
drususallee 1
39100 BOZEN / BOLZANO [IT]
- up to 20 participants
- costs: **10,00 €**
incl coffee break



please feel free to **contact:**
eveline.wandl-vogt@oeaw.ac.at