

WG3 Innovative e-dictionaries

Proposal for work plan

The original proposal for COST Action IS1305 specifies a number of topics that will be dealt with within the Action in Working Group 3. These are:

- Description of the workflow for corpus-based lexicography
- Overview of existing software needed in this workflow
- Dictionary Writing Systems (and Corpus Query Systems)
- Analysis of the possible impact of automatic acquisition of lexical data (distributional thesauri etc.)
- Analysis of the interface between dictionaries and computational lexica (cf. wordnets) and syntactically and semantically annotated corpora (Framenet, Semcor, Senseval)
- Investigation of possible use of dictionary content for computational linguistic applications

We propose to organise the meetings of WG3 around these topics dealing ideally with one theme/topic per meeting. We would like to identify a key person for each meeting who will take the lead for that particular meeting (thus sharing the work load).

July 2014: **Workflow of corpus-based lexicography; Software to support lexicographical workflow (also backup, version control etc.)** (Tiberius)

Result: better understanding of the workflow which results in better planning of future projects

January 2015: **Software to support lexicographical workflow: DWS and CQS** (Krek)

Result: a description of DWSs and in particular the newly developed (web) applications for querying corpora

July 2015: **Automatic acquisition of lexical data and its impact (what works, what doesn't work – example sentences, collocations, neologisms, definitions, word senses)**

Result: exploring the possibility of automation of particular tasks within corpus-based lexicography as a support to lexicographers / lexicographical workflow

January 2016: **Between Corpora and Dictionaries – analysis of the interface between dictionaries and computational lexica and (syntactically and semantically annotated) corpora**

Result: exploring the possibility of collecting lexically and semantically organized data in a completely automated process where the data could be used for immediate visualization for human users interested in lexical behaviour of words

July 2016: **The use of lexicographical data in computational linguistics – investigation of possible use of dictionary content for computational linguistic applications**

Result: better understanding of the need of computational linguistic community for lexicographically organized data and vice versa

COST Action IS1305

WG3 Innovative e-dictionaries

DATE: 16 JANUARY 2014, 11H15-13H00

PLACE: LEIDEN, THE NETHERLANDS

PROGRAMME

11H15-11H35 PRACTICALITIES

11H35-12H15 WORK PLAN & TIME-TABLE

12H15-12H40 DEFINITION OF TASKS FOR NEXT MEETING

12H40-12H50 SUGGESTIONS FOR AN APPLICATION FOR A 5-DAY MEETING IN THE LORENTZ CENTER (LEIDEN) IN 2016

12H50-13H00 AOB AND CLOSING