

## Visualization Design for a Web Interface to the Large-Scale Linked Lexical Resource UBY

We present the results of a collaboration of visualization experts and computational linguists which aimed at the re-design of the visualization component in the Web user interface (Web UI) to the large-scale linked lexical resource UBY. UBY combines a wide range of information from expert-constructed (e.g., WordNet, FrameNet, VerbNet) and collaboratively constructed (e.g., Wiktionary, Wikipedia) resources for English and German, see <https://www.ukp.tu-darmstadt.de/uby>. All resources contained in UBY distinguish not only different words but also their senses. A distinguishing feature of UBY is that the different resources are aligned to each other at the word sense level, i.e. there are links connecting equivalent word senses from different resources in UBY. For senses that are linked, information from the aligned resources can be accessed and the resulting enriched sense representations can be used to enhance the performance of Natural Language Processing tasks.

Targeted user groups of the UBY Web UI are researchers in the field of Natural Language Processing and in the Digital Humanities (e.g., lexicographers, linguists). In the context of exploring the usually large number of senses for an arbitrary search word, the UBY Web UI should support these user groups in assessing the added value of sense links for particular applications. It is important to emphasize that this is an open research question for most applications.

We will present the results of our detailed requirements analysis that revealed a number of central requirements a visualization of all the senses for a given search word and the links between them must meet in order to be useful for this purpose. The existing implementation of the visualization component in the UBY Web UI does not meet these requirements. In particular, it does not scale to words with many senses, such as “run” having ca. 200 senses in UBY, see figure 1.

In our talk, we will present the revised visualization component, see figure 2, which is the result of an iterative and highly collaborative process consisting of several stages of designing visualization mockups, and rigorously analyzing and discussing them in light of the requirements. The final visualization groups senses within the same cluster and encloses them in a circle. Within each cluster the senses are ordered according to resources (distinct colors are used for color-encoding the resource membership of a sense). This visualization was convincing for the computational linguists, because it is intuitive to read and far less complex than the original visualization.

Finally, we will summarize the lessons learned in this interdisciplinary collaboration between visualization experts and computational linguists. This collaboration also reminded us again that it takes a considerable amount of time and several iterations until both sides gain a sufficient amount of knowledge about both the capabilities of visual analysis and the specific task considered.

