

WordNet-Inspired Terminological Resources for Bio-NLP

Elena Beisswanger, Michael Poprat, Udo Hahn
Jena University Language & Information Engineering (JULIE) Lab,
Jena, Germany

WordNet is currently the most widely used lexicon resource for general English language. We here argue in favor of a similar lexical resource for biomedicine, BioWordNet, to extend the virtues of WordNet to this sublanguage domain. We present a simple approach to semi-automatically build up such a resource. It crucially builds on the conversion of structured domain knowledge taken from the Open Biomedical Ontologies (OBO) and the subsequent mapping of this library to WordNet's lexicographic file format. We report on the shortcomings of the original WordNet format that hampered this approach. Subsequently we propose alternative strategies to construct a BioWordNet in a more up-to-date representation such as the Web Ontology Language (OWL). Finally we point out the steps ahead when building up such a domain-specific resource from various sources and discuss the general problems that need to be solved when integrating the BioWordNet into the original WordNet.