Conference Abstract

Using Semantics for morphological Descriptions in Morph•D•Base

Christian Köhler[‡], Roman Baum[‡], Peter Grobe[§], Sandra Meid[‡], Björn Quast[§], Lars Vogt[‡]

‡ Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany
§ Zoological Research Museum Alexander Koenig, Bonn, Germany

Abstract

Providing data in a semantically structured format has become the gold standard in data science. However, a significant amount of data is still provided as unstructured text - either because it is legacy data or because adequate tools for storing and disseminating data in a semantically structured format are still missing. We have developed a description module for Morph·D·Base, a semantic knowledge base for taxonomic and morphologic data, that enables users to generate highly standardized and formalized descriptions of anatomical entities using free text and ontology-based descriptions. The main organizational backbone of a description in Morph·D·Base is a partonomy, to which the user adds all the anatomical entities of the specimen that they want to describe. Each element of this partonomy is an instance of an ontology class and can be further described in two different ways:

1. as semantically enriched free-text description that is annotated with terms from ontologies, and
2. semantically through defined input forms with a wide range of ontology-terms to choose from.

To facilitate the integration of the free text into a semantic context, text can be automatically annotated using jAnnotator, a javascript library that uses about 700 ontologies with more than 8.5 million classes of the National Center for Biomedical Ontology (NCBO) bioportal.

Users get to choose from suggested class definitions and link them to terms in the text, resulting in a semantic markup of the text. This markup may also include labels of elements that the user already added to the partonomy. Anatomical entities marked in the text can be added to the partonomy as new elements that can subsequently be described semantically using the input forms. Each free text together with its semantic annotations is stored following the W3C Web Annotation Data Model standard (https://www.w3.org/TR/annotation-model). The whole description with the annotated free text and the formalized semantic descriptions for each element of the partonomy are saved in the tuplestore of Morph·D·Base.

The demonstration is targeted at developers and users of data portals and will give an insight to the semantic Morph·D·Base knowledge base (https://proto.morphdbase.de) and jAnnotator (http://git.morphdbase.de/christian/jAnnotator).

Keywords

ontology, semantic annotation, javascript, Linked Open Data, Morph·D·Base, knowledge base, Semantic Instance Anatomy

Presenting author

Christian Köhler

Funding program

Deutsche Forschungsgemeinschaft (DFG): http://gepris.dfg.de/gepris/projekt/248394582?language=en

Grant title

Digitisation / Cataloguing of non-textual objects: eScience-compliant standards for morphology