Conference Abstract

# Using Named Graphs and Knowledge Graph Template Patterns for Efficiently Organizing FAIR Anatomy Data and Metadata

Lars Vogt[‡], Roman Baum[‡]

‡ Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

## Abstract

Currently, morphological data and metadata are still mostly published as unstructured free texts, which lack semantic transparency, cannot be parsed by computers, and do not comply with the FAIR (Findable, Accessible, Interoperable, Reusable; Wilkinson et al. (2016) data principles, thus hampering their reuse by non-experts and their integration across many fields in the life sciences. With an ever-increasing amount of available ontologies and the development of adequate semantic technology, however, a solution to this problem becomes available.

Instead of free text descriptions, morphological data and metadata can be recorded, stored, and communicated through the Web in the form of Resource Description Framework (RDF) triple statements that use the *'Subject – Property – Object'* syntax of RDF and URIs of ontology classes and properties as well as URIs for individual entities as terminology. Since a given URI can take the *'Subject'* position in one and the *'Object'* position in another RDF statement, several triples can be linked to form a highly formalized and structured directed graph (semantic graph).

After introducing an instance-based approach of recording morphological descriptions and their accompanying metadata as semantic knowledge graphs (i.e. *Anatomy Knowledge*

*Graphs*), we propose a knowledge graph template pattern for each type of anatomical observation and a pattern for documenting metadata. The use of template patterns for knowledge graphs provides *Interoperability* and *Reusability* of comparable anatomical observations and of their accompanying metadata and a means to meaningfully visualize information contained in semantic graphs in a user-friendly HTML representation. Stored in a tuple store, *Anatomy Knowledge Graphs* become *Findable* and *Accessible* through the store's SPARQL endpoint. As a consequence, anatomy data and metadata documented as *Anatomy Knowledge Graphs* in a tuple store are FAIR.
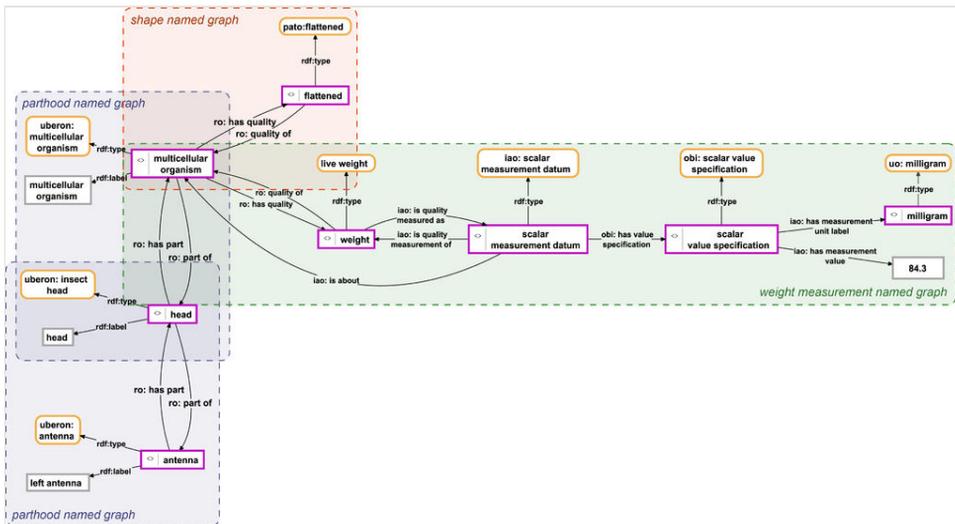


Figure 1.

**Anatomy Knowledge Graph**

The semantic knowledge graph shows the description of a multicellular organism. It consists of instances (purple-bordered boxes), each of which instantiates a specific ontology class (yellow-bordered boxes with rounded corners) through the *type* property. All instances referring to anatomical entities possess a human-readable label (grey-bordered box connected through the *label* property) and are connected with one another via parthood relations, forming a partonomy indicating that the **multicellular organism** instance possesses an **insect head** instance that, in turn, possesses an **antenna** instance. The **multicellular organism** instance is further described to have a **flattened** shape and a measured **live weight** of 84.3 **milligram**s. The semantic graph is organized and fragmented into different sub-graphs, each of which is contained in its own named graph (dashed-bordered colored boxes). Each sub-graph contains information relating to a specific perceptual question that can only be answered empirically. In other words, each named graph contains a separate empirical observation. The *Semantic Instance Anatomy* is the union of all the named graphs.

Finally, we suggest a general scheme of how to efficiently organize *Anatomy Knowledge Graphs* in a tuple store framework based on instances of named graphs, with each individual named graph instantiating an ontology class that relates to a particular type of observation (e.g., weight measurement named graph class). A named graph is a fourth

element in an RDF statement (*'Subject – Property – Object – Named Graph'*), turning the triple into a quadruple. All RDF statements that share the same URI in the *'Named Graph'* position belong to the same named graph. The use of named graph resources allows meaningful fragmentation of the contents of an *Anatomy Knowledge Graph* (Fig. 1), which in turn enables subsequent specification of all kinds of data views for managing and accessing morphological data and metadata. This scheme has been implemented in the description module of the prototype for [semantic Morph·D·Base](#).

## Keywords

Anatomy Knowledge Graph; anatomy; knowledge graph; named graph; morphological description; metadata; FAIR data principle; semantically structured data; ontology

## Presenting author

Lars Vogt

## Presented at

Biodiversity_Next 2019

## Acknowledgements

## Grant title

eScience-Compliant Standards for Morphology

## References

- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, 't Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M,

Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3 https://doi.org/10.1038/sdata.2016.18