

Comprehensive data annotation and findable data: Mapping odML to RDF

Michael Sonntag1, Yaroslav Shalivskyy1, Achilleas Koutsou1, Christian Garbers1, Jan Grewe2, Thomas Wachtler1





¹German Neuroinformatics Node, Department Biologie II, Ludwig-Maximilians-Universität München, Germany; ²Institut für Neurobiologie, Universität Tübingen, Germany



Overview

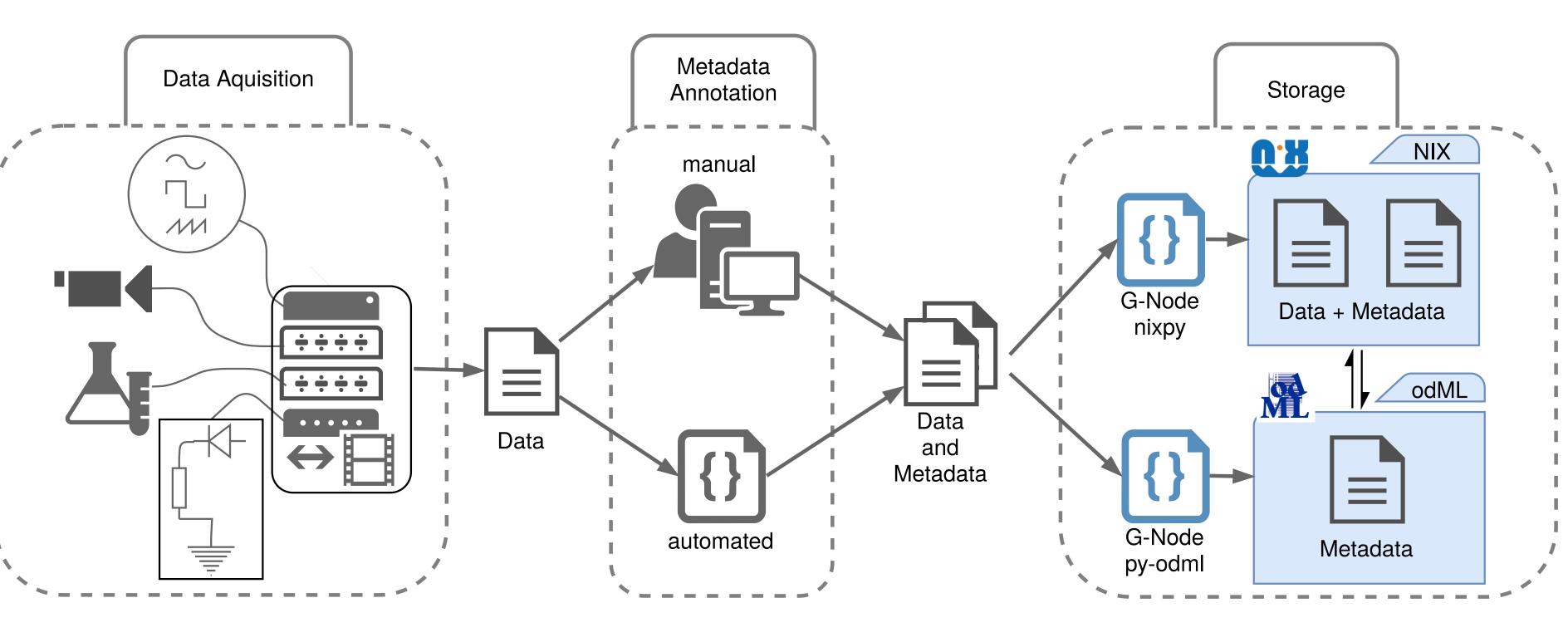
Annotation of research data with metadata is vital to provide context for analysis and data re-use. The odML[1] format enables collecting metadata from different sources in an organized, flexible, human and machine-readable fashion[2] and is easy to use for the scientist. odML specifies the format, not the content, so any metadata necessary to describe a given dataset can be stored.

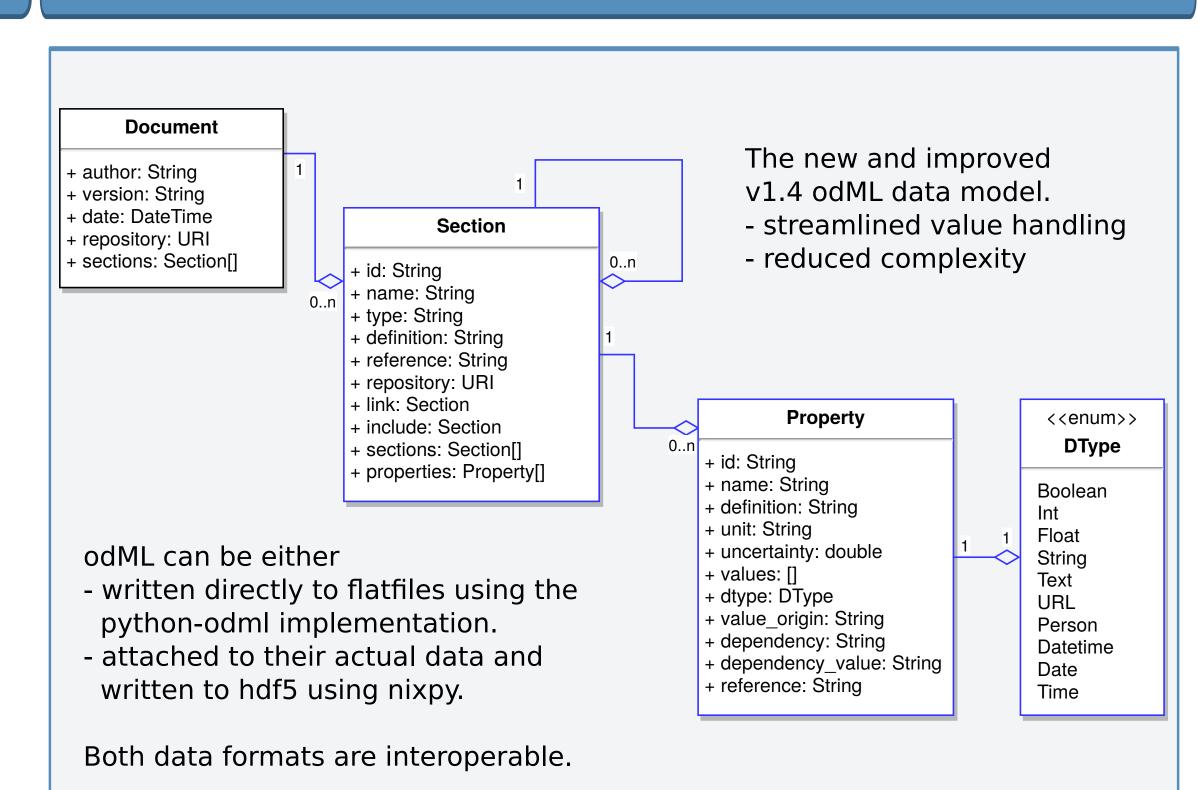
Building on the odML format we present an approach utilizing Semantic Web technologies to effortlessly make even diverse metadata interoperable, findable and accessible according to the FAIR principles[3]. With a small set of terms derived from the odML format we defined a mapping from the odML data model to a general repesentation in RDF[4] and developed a straightforward conversion pipeline.

Thus, metadata collected via the convenient odML format can be fed to a single local or distributed searchable RDF graph. Taking advantage of the powerful OWL language[5], each distinct set of metadata can be subclassed further to the benefit of maintaining the original relations without losing the common structure, achieving interoperability without sacrificing findability. To enable easy access to large collections of metadata, we developed a custom "fuzzy search" feature and further introduce an augmented SPARQL server based on the renown Apache Jena Fuseki[6] that offers a convenient search interface for the scientist by using the powerful SPARQL query language.

A data/metadata acquisition workflow featuring odML

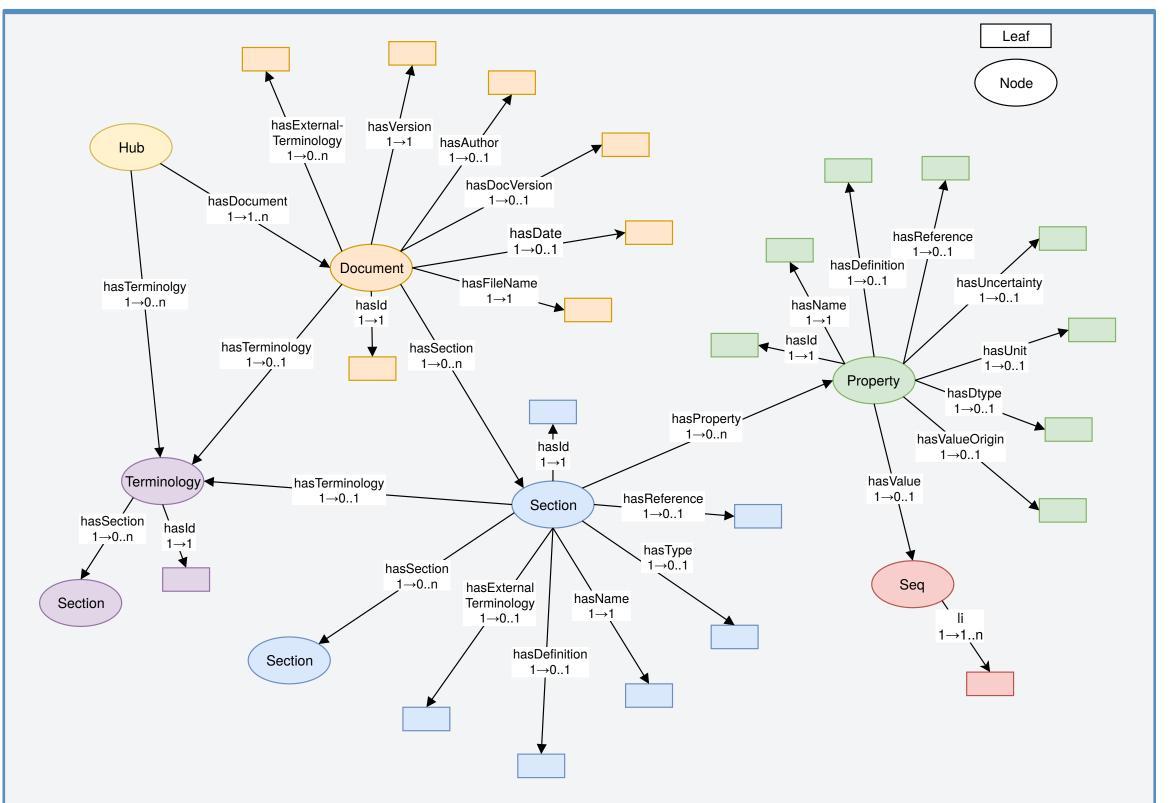
odML data model

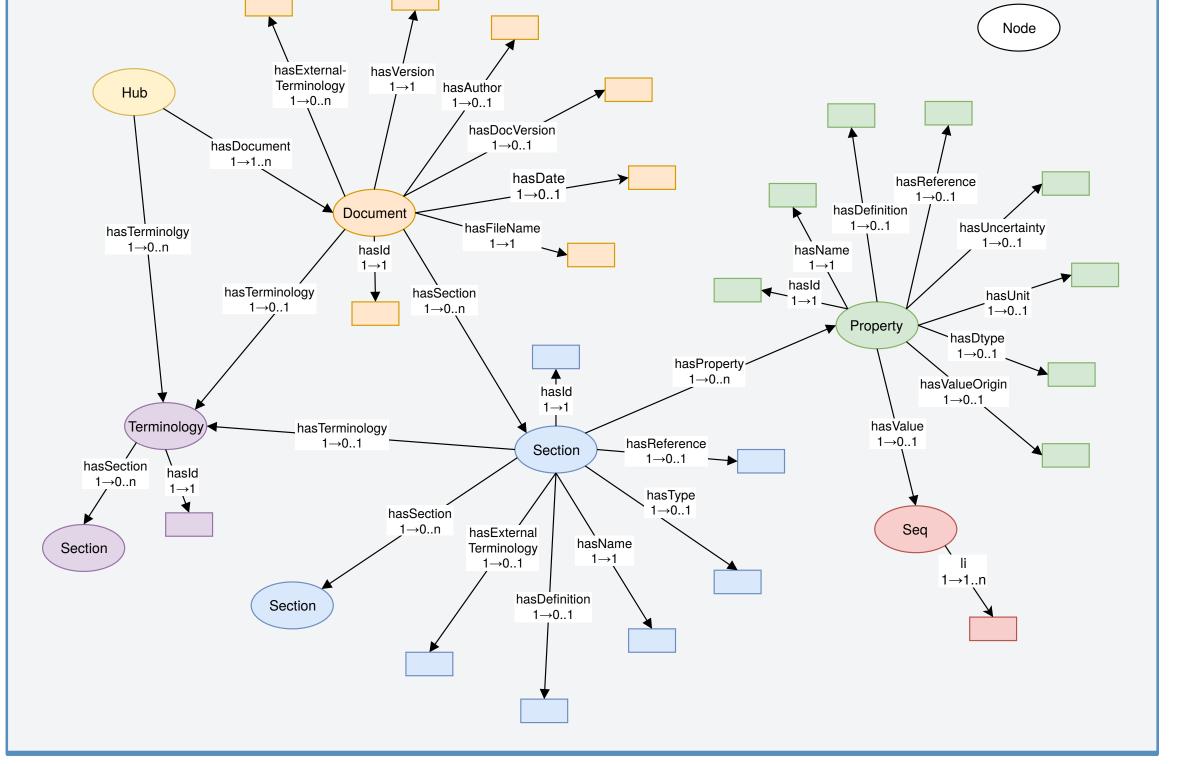




odML flavored, basic RDF

Opening Metadata to the Semantic Web via odML flavored RDF

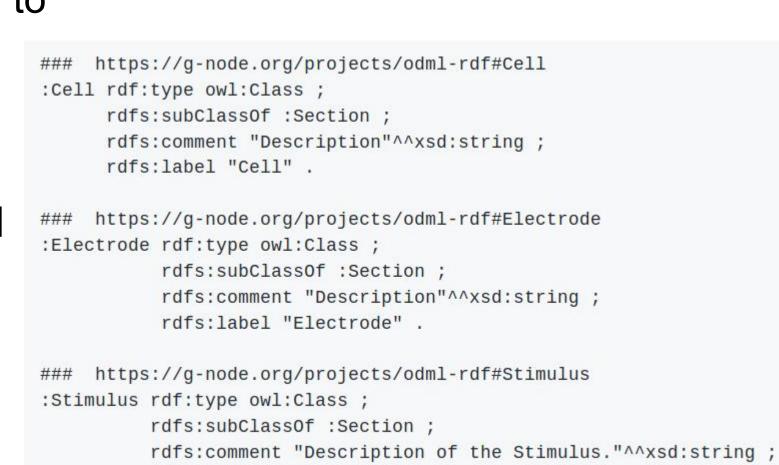


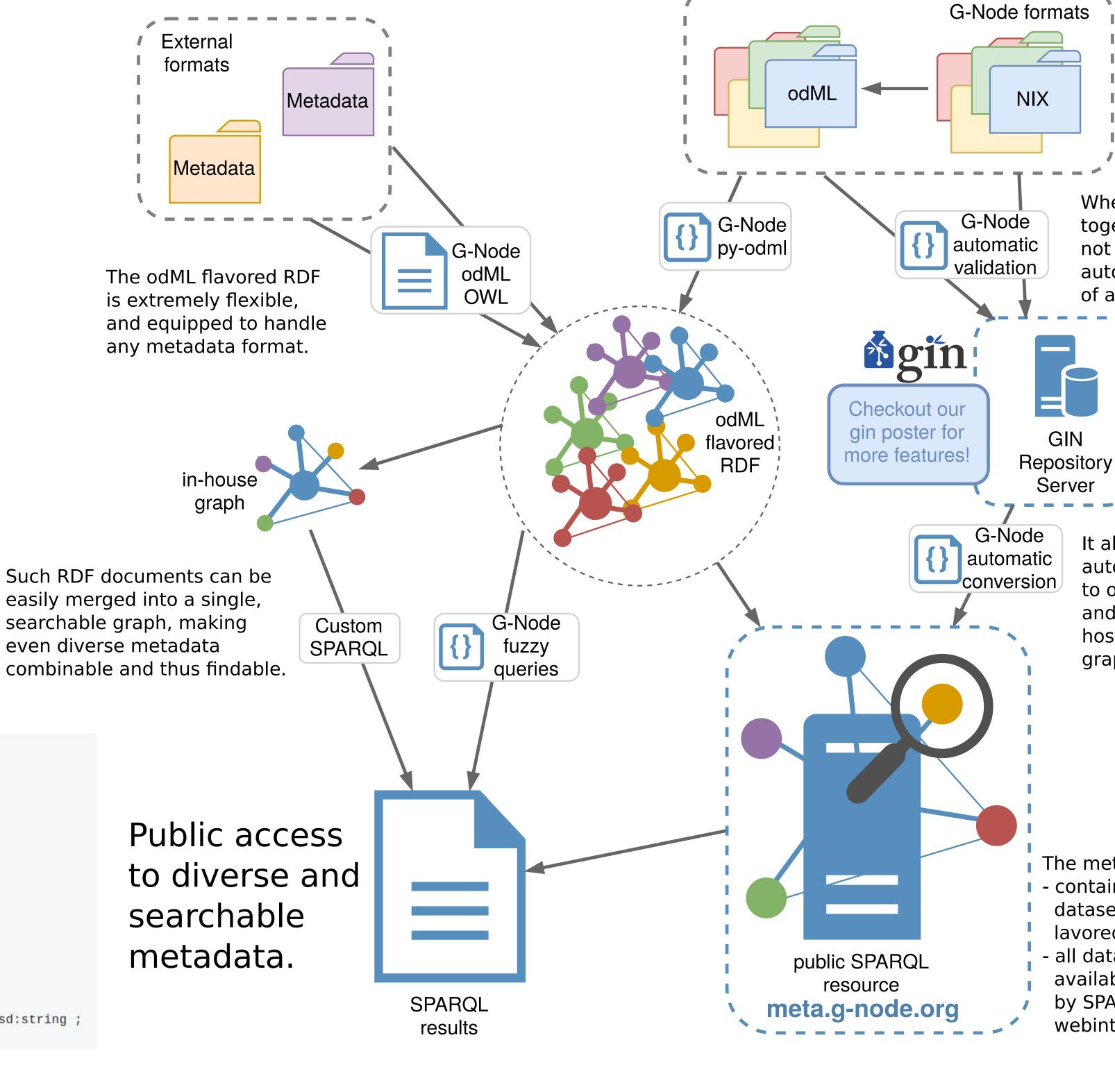


Using OWL to finegrain RDF metadata

Extend the basic odML OWL ontology by subclassing the basic odML flavored RDF to

- retain your own metadata structure and terms even in RDF.
- enable more finegrained SPARQL queries for your needs while enabling general searches through still valid basic odML RDF terms.





It also provides optional automatic conversion to odML flavored RDF and upload to a G-Node hosted public metadata graph database.

When using nix or odML

together with gin, this

automatic validation

not only provides

of any file.

The meta.g-node.org server - contains diverse metadata datasets in the odML f lavored RDF format. - all datasets are publicly available and searchable by SPARQL via API and webinterface.

Resources and references



Find the odml, nix and gin projects at

rdfs:label "Stimulus"

http://meta.g-node.org https://github.com/G-Node/python-odml https://github.com/G-Node/nixpy https://gin.g-node.org

Find out about more g-node projects at https://g-node.github.io

Contact: dev@gnode-org

References

[1] Grewe et al (2011); doi:10.3389/fninf.2011.00016

[2] Zehl et al (2016); doi:10.3389/fninf.2016.00026 [3] Wilkinson et al (2016); doi:10.1038/sdata.2016.18

[4] https://www.w3.org/RDF/

[5] https://www.w3.org/TR/owl-features/ [6] https://jena.apache.org/documentation/fuseki2/







Developed at the LMU Munich

Supported by BMBF grants 01GQ1302 and 01GQ1509