Comprehensive data annotation and findable data: Mapping odML to RDF

Michael Sonntag[1], Yaroslav Shalivskyy[1], Achilleas Koutsou[1], Christian Garbers[1], Jan Grewe[2], Thomas Wachtler[1]

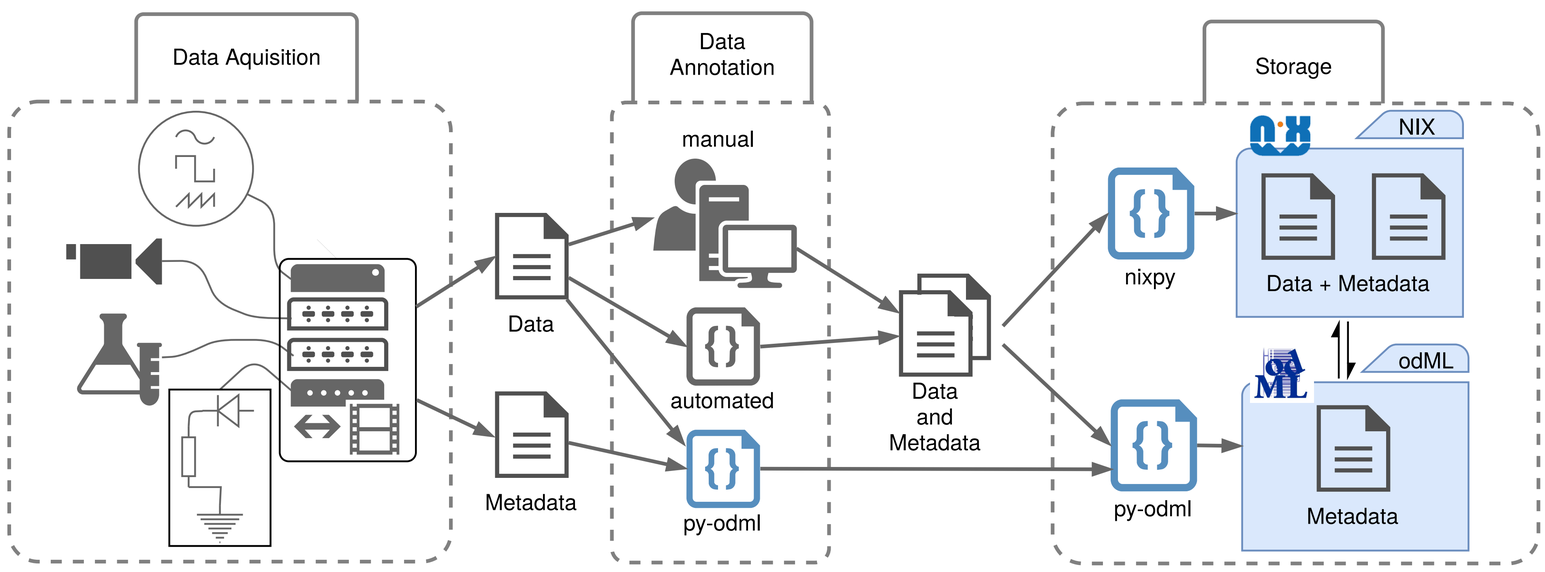
[1]German Neuroinformatics Node, Department Biologie II, Ludwig-Maximilians-Universität München, Germany;
[2]Institut für Neurobiologie, Universität Tübingen, Germany

Overview

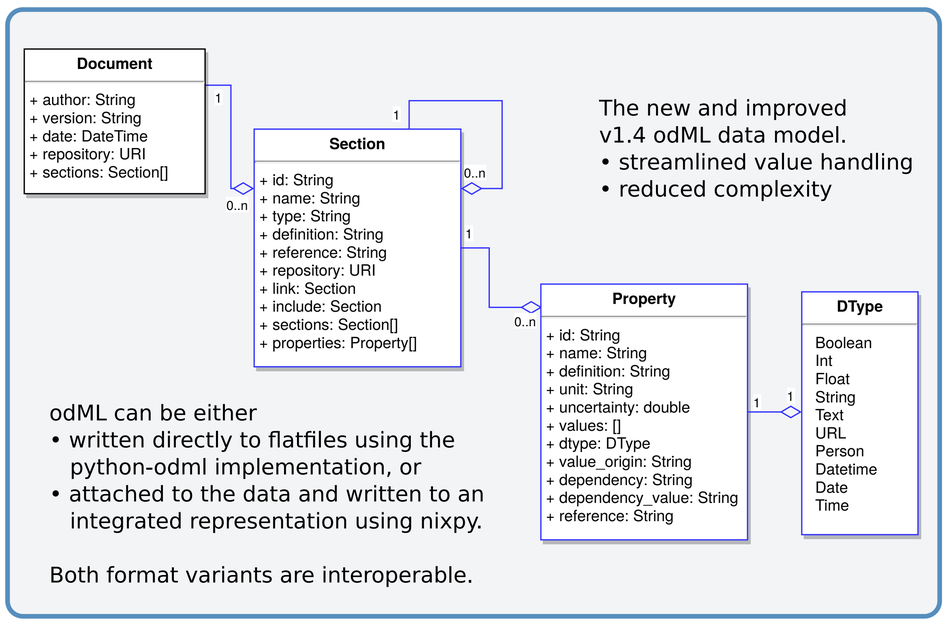
Annotation of research data with metadata is vital to provide context for analysis and data re-use. The odML[1] format enables collecting metadata from different sources in an organized, flexible, human and machine-readable fashion[2] and is easy to use for the scientist. odML specifies the format, not 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 Apache Jena Fuseki[6] that offers a convenient search interface for the scientist by using the powerful SPARQL query language.

Data / Metadata acquisition workflow using odML



odML data model

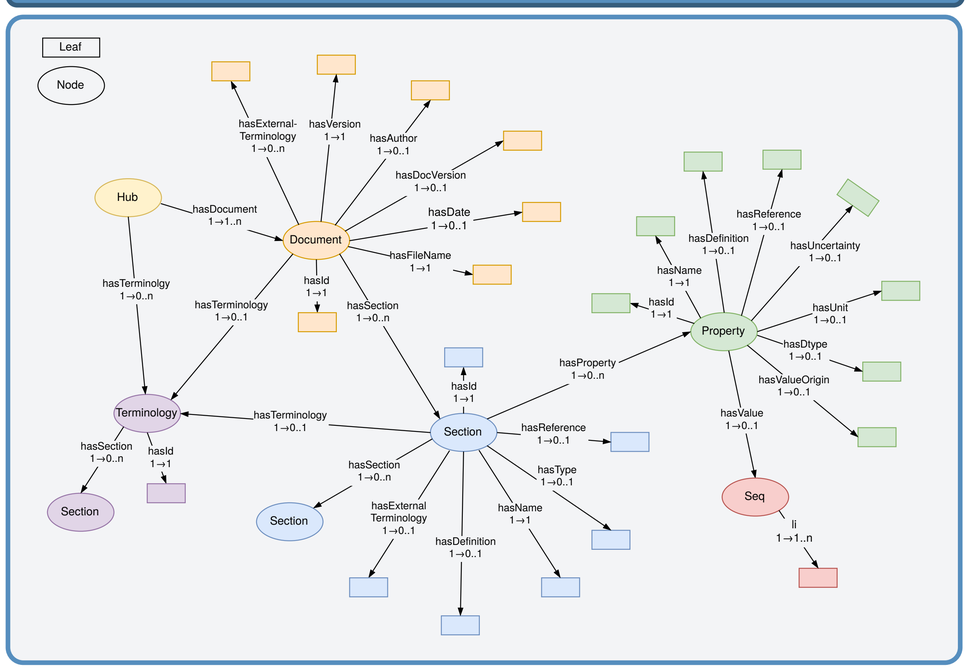


The new and improved v1.4 odML data model.
- streamlined value handling
- reduced complexity

odML can be either
- written directly to flatfiles using the python-odml implementation, or
- attached to the data and written to an integrated representation using nixpy.

Both format variants are interoperable.

RDF representation of odML data model



Using OWL to fine-tune RDF Metadata

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

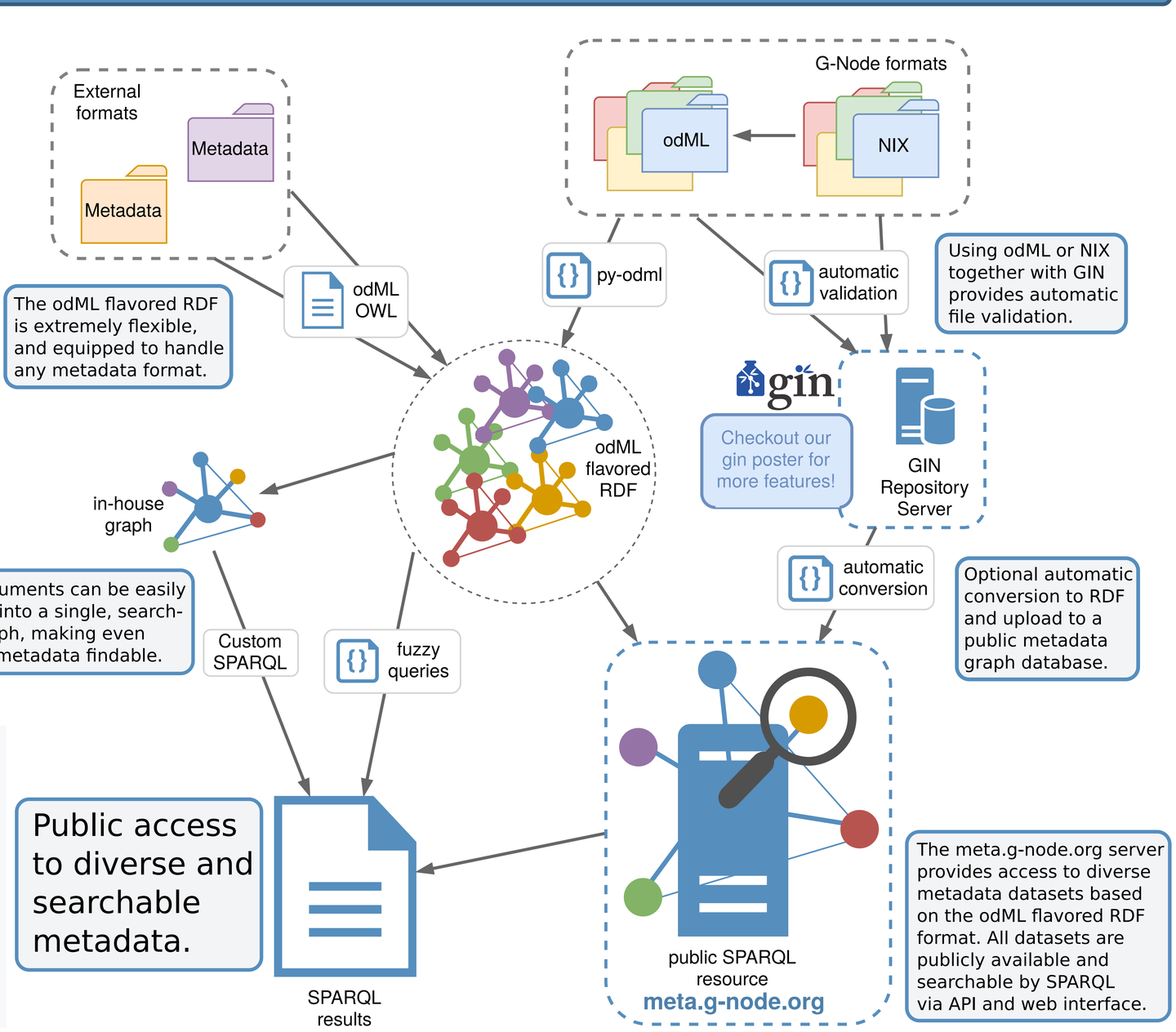
- retain your own metadata structure and terms even in RDF.
- enable more specific SPARQL queries for your needs while enabling general searches through odML terms.

```
###  https://g-node.org/projects/odml-rdf#Cell
:Cell rdf:type owl:Class ;
      rdfs:subClassOf :Section ;
      rdfs:comment "Description"^^xsd:string ;
      rdfs:label "Cell" .

###  https://g-node.org/projects/odml-rdf#Electrode
:Electrode rdf:type owl:Class ;
           rdfs:subClassOf :Section ;
           rdfs:comment "Description"^^xsd:string ;
           rdfs:label "Electrode" .

###  https://g-node.org/projects/odml-rdf#Stimulus
:Stimulus rdf:type owl:Class ;
          rdfs:subClassOf :Section ;
          rdfs:comment "Description of the Stimulus."^^xsd:string ;
          rdfs:label "Stimulus" .
```

Opening Metadata to the Semantic Web via odML



Resources and references

Find the odML, NIX and GIN projects at

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@g-node.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/

Poster presented at
INCF Neuroinformatics 2018
Montréal, Canada
doi:10.12751/incf.ni2018.0110

Hosted by LMU Munich

Supported by BMBF grants 01GQ1302 and 01GQ1509