

Conceptualising RDFBones as an RDF ontology and application

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

Felix Engel and Stefan Schlager

Biologische Anthropologie Freiburg

Workshop “Digital Standards for Research Data from Human Skeletal Collections”,
06.10.2016

Human Skeletal Collections

Collections surveyed for this project as examples



UNI
FREIBURG

■ Historical research collections

- Alexander Ecker Collection, Freiburg
- Rudolf Virchow Collection, Berlin
- Blumenbach Collection, Göttingen

■ Active research collections (various contexts)

- Osteological Collection, Tübingen

■ Bioarchaeological archives

- State Collection of Anthropology and Palaeoanatomy Munich
- Archaeological Unit Baden-Württemberg, Constance

■ Museums (various contexts)

- Natural History Museum Basel
- (Natural History Museum Vienna)

■ Forensic facilities

Materials to be Modelled

Human anatomy



UNI
FREIBURG

Human remains

- Skeletal material
 - Skeletons
 - Commingled remains
 - Material from stratigraphical units
 - Forensic cases
- Mummified remains
- Wet specimen

Models

- Anatomical
- Pathological

Replica

- Humans
- Human remains

Materials to be Modelled

Material objects



UNI
FREIBURG

Objects

- Archaeological finds
 - Artefacts
 - Biological/geological samples
- Ethnological objects

Replica

- Archaeological artefacts
- Ethnological Objects

Materials to be Modelled

Documents



Unpublished

- Earlier collection inventories
- Reports
 - Research projects
 - Expeditions
 - Archaeological excavations
 - Forensic cases
- Personal records
 - Diaries
 - Laboratory notebooks
 - Notes, sketches
- Correspondence
 - Personal letters
 - Business correspondence

Art

- Drawings
- Paintings

Published

- Literature
 - Monographs
 - Text books
 - Academic journals
 - Manuals
 - Catalogues
- Maps

Materials to be Modelled

Referenced vs. unreferenced materials

It is assumed that each collection has an active inventory that indexes all objects comprised in the collection.

Referenced materials

Materials listed in the collection inventory.

Unreferenced materials

Materials that are kept with the collection because they are relevant to the collection objects but which do not appear in the collection's inventory.

Processes to be Modelled

Collection history

- Acquisitions, disposals of objects
- Curation periods
- Institutional changes
- Changes of locality
- Decisions
(e. g. restitution processes)

Research

- Projects
- Investigations
- Sampling

Actors to be Modelled

Persons

- Users
- Researchers
- Curators
- Contributors
- Identities (e.g. of skeletal individuals)

Institutions

- Keepers of the collection
- Research facilities
- Interest groups
- Service providers

Additional Information to be Modelled



UNI
FREIBURG

Content-oriented

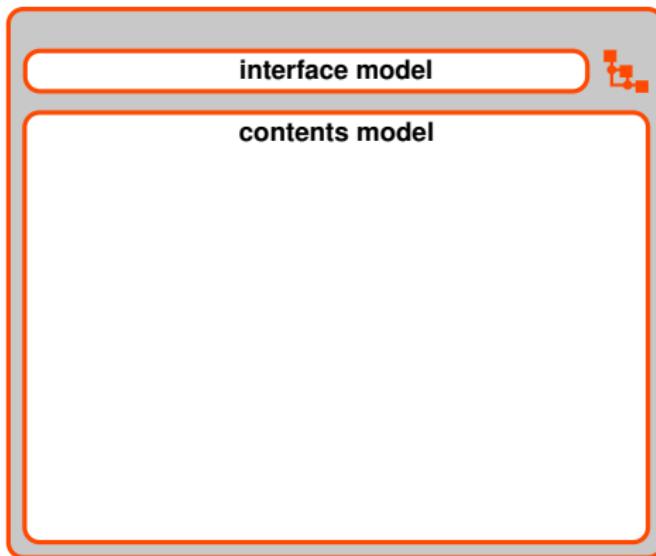
- Workflow-related information
- Notes
- Restrictions on access to material
- Directives for material preservation

Display-oriented

- Workflow-related information
- Information on form structure

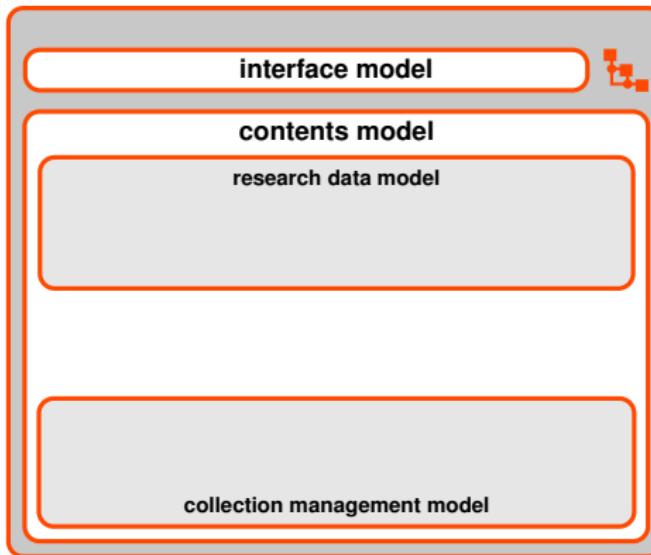
Data Model

Contents and Interface



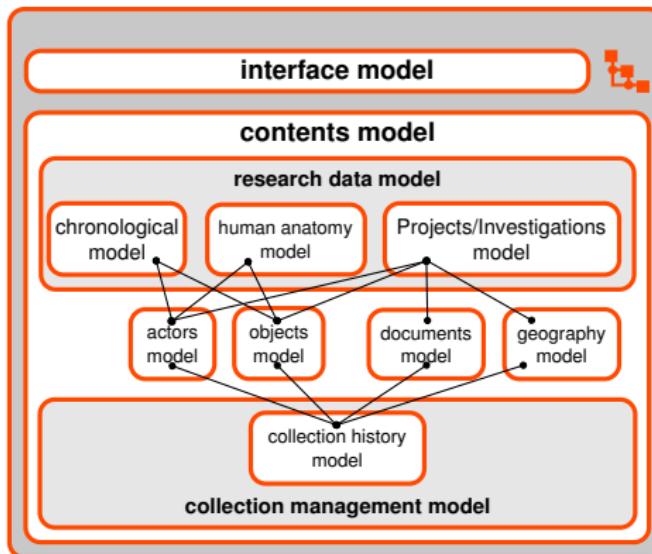
Data Model

Research Data and Collection Management



Data Model

Modules

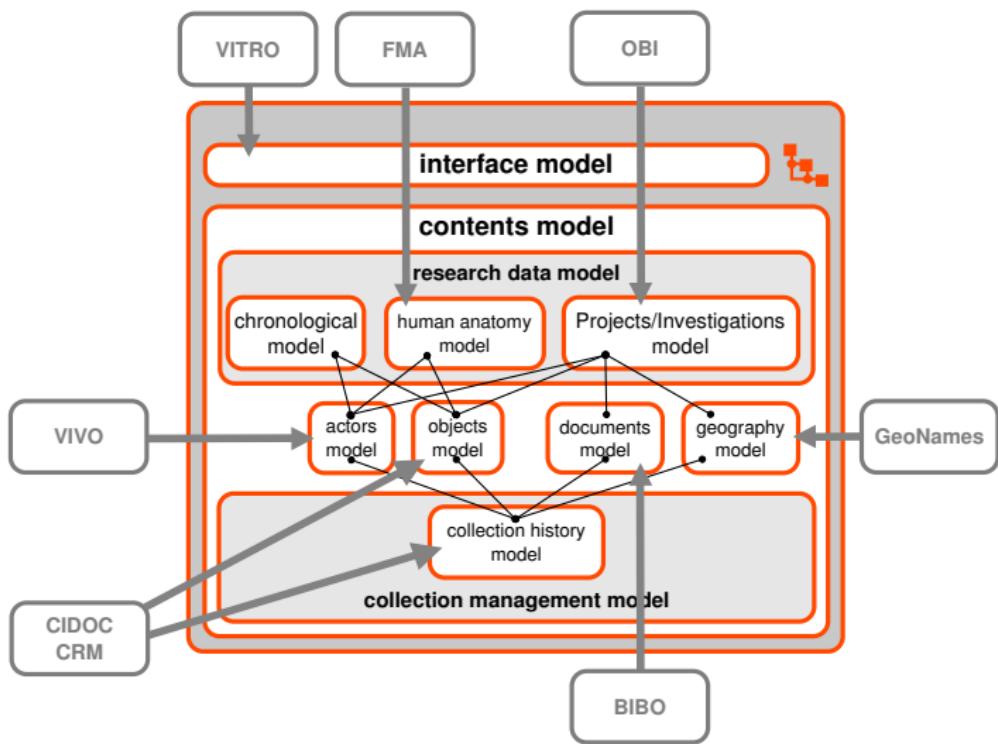


Data Model

External Ontologies



UNI
FREIBURG

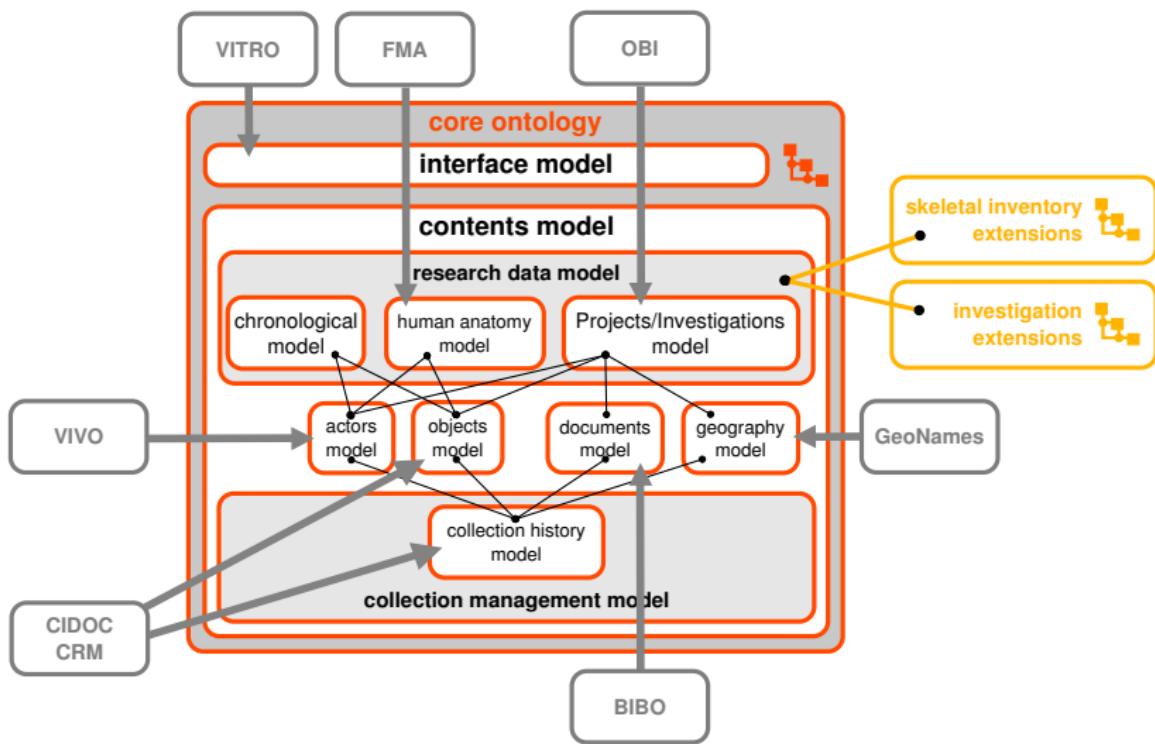


Data Model

Core Ontology and Extensions

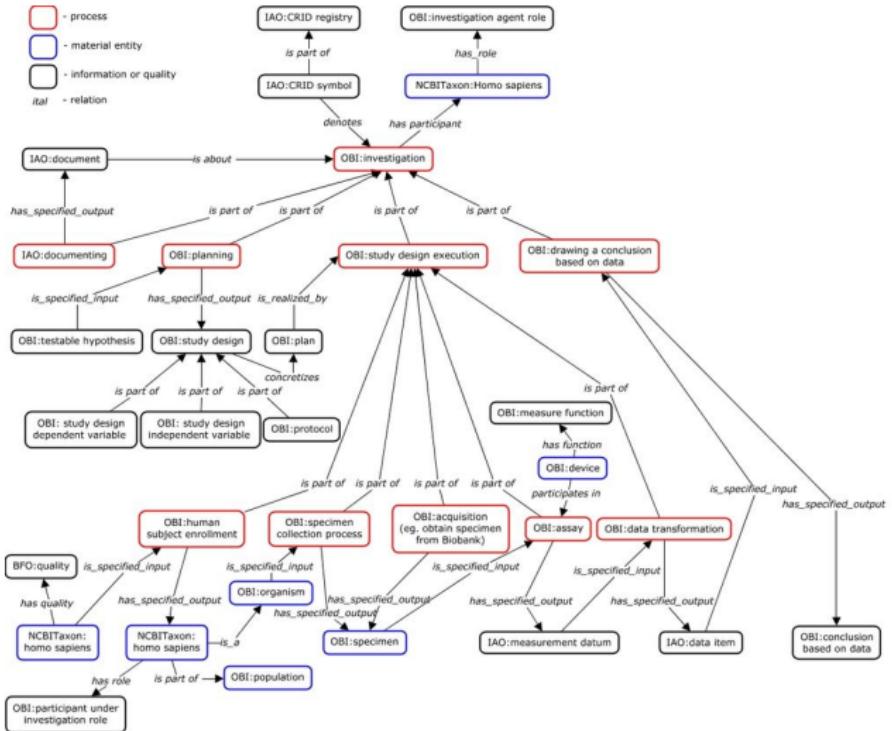


UNI
FREIBURG



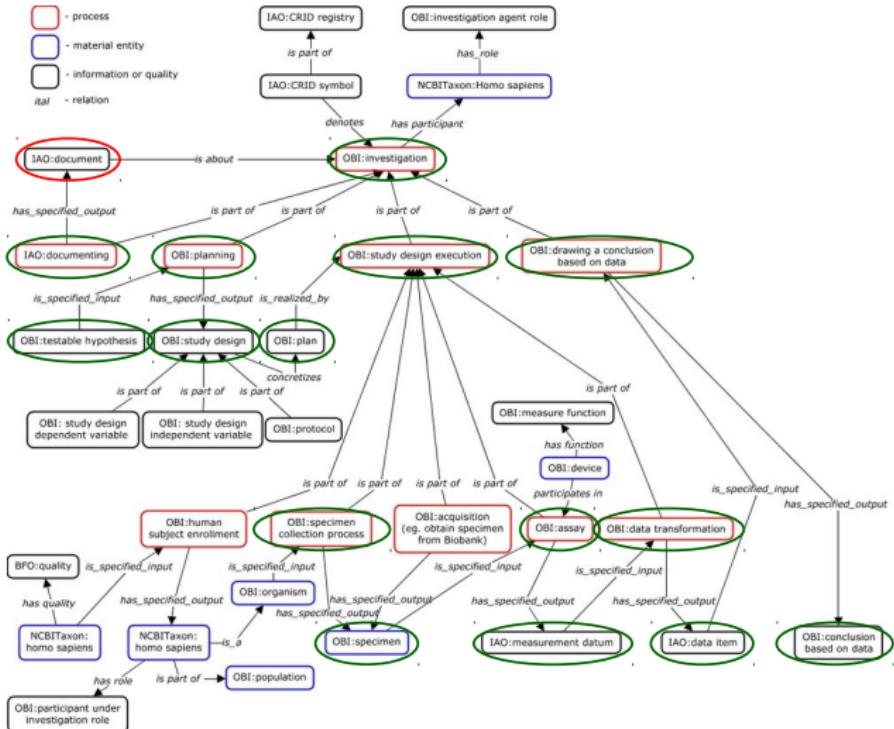
Ontology for Biomedical Investigations

An Example for an External Ontology



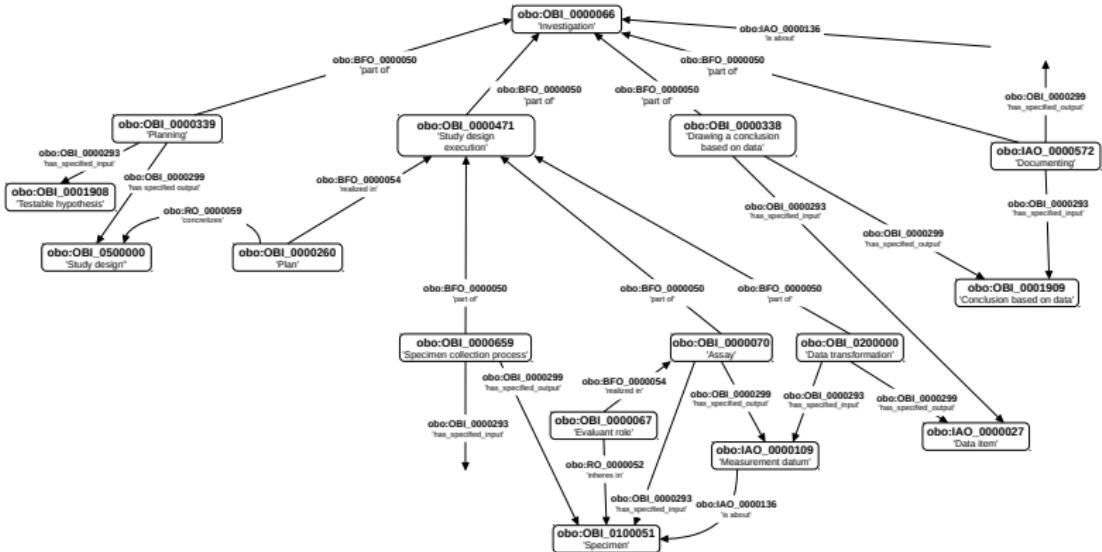
Ontology for Biomedical Investigations

Class Not Used in RFBones



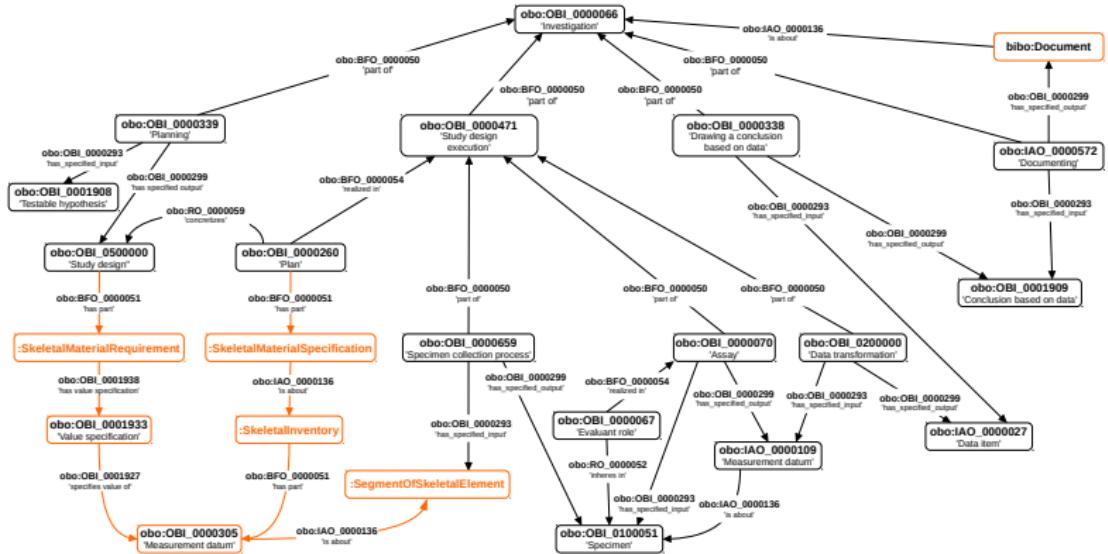
Modelling Investigation Data

Classes Selected from OBI Documentation



Modelling Investigation Data

Classes Selected from OBI Documentation



Core Ontology and Extensions



Problem of Duplicate Creation

Referenced materials

Separation of data input and reference

- > primary directories
- > primary inventories

Non-referenced materials

Creation of inventory upon data entry

External materials

Duplicates likely, use owl:sameAs to eradicate.

The Ontologist's Workshop

RDF-Editor Protégé and Text Editor

Screenshot of the Protégé RDF-Editor showing the Class hierarchy for 'Investigation' and its sub-classes.

Class hierarchy (Class hierarchy inferred):

- Investigation**
 - Event
 - Event Series
 - Planning Process'
 - Phase
 - Planned Process'
 - assay
 - assessing information'
 - documenting'
 - drawing a conclusion based on evidence'
 - Learning'
 - Research Project'
 - Specimen collection process'
 - study design execution'
 - study design planning'
 - Topic Weight Measurement Project'
 - planning'
 - Presenting Process'
 - domain'

Annotations:

- label** [language: en]
- investigation**

definition [language: en]

a planned process that consists of parts: planning, study design execution, documentation and which produce conclusions.

transient note [type: ext-string]

Could add specific objective specification

curator note [language: en]

Following OBI call November 2012,26th it was decided there was no need for adding 'achieves objective of drawing conclusion' as existing relations were providing equivalent utility; this note closes the issue and validates the class definition to be part of the OBI core editor = PRS

Description:

investigation

Is-a:

- Has OBI class
- Planning
- and
- (has specified out-of: some 'Study Design')
- has part some 'drawing a conclusion based on data'
- has part some 'study design execution'
- has part some 'documenting'
- Planned Process'

General class axioms:

SubClass Of (Anonymous Axiom)

Instances:

Target for key

Depict with

Depict Union of

To use the reasoner click Reasoner > Start reasoner Show inferences

Screenshot of the Emacs text editor showing the XML code for the OBI-RDBBonesSubset.owl ontology.

```

<owl:Restriction>
  <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/BFO_0000051"/>
  <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/IAO_00000572"/>
</owl:Restriction>
<rdfs:subClassOf>
<owl:Restriction>
  <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/BFO_0000051"/>
  <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/IAI_00000338"/>
</owl:Restriction>
<rdfs:subClassOf>
<owl:Restriction>
  <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/BFO_0000051"/>
  <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/IAI_00000471"/>
</owl:Restriction>
<rdfs:subClassOf>
<owl:Restriction>
  <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/BFO_0000051"/>
  <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/IAI_00000299"/>
</owl:Restriction>
<owl:intersectionOf rdf:type="Collection">
  <rdf:Description rdf:about="http://purl.oblibrary.org/obo/OBI_00000339"/>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/OBI_00000299"/>
    <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/OBI_0594"/>
  </owl:Restriction>
</owl:intersectionOf>
</owl:Restriction>
<owl:Restriction>
  <owl:onPropertiesFrom>
    <owl:IntersectionOf rdf:type="Collection">
      <rdf:Description rdf:about="http://purl.oblibrary.org/obo/OBI_00000339"/>
      <owl:Restriction>
        <owl:onProperty rdf:resource="http://purl.oblibrary.org/obo/OBI_00000299"/>
        <owl:someValuesFrom rdf:resource="http://purl.oblibrary.org/obo/OBI_0594"/>
      </owl:Restriction>
    </owl:IntersectionOf>
  </owl:onPropertiesFrom>
</owl:Restriction>
<owl:intersectionOf rdf:type="Collection">
  <rdf:Description rdf:about="http://www.w3.org/2001/XMLSchema#string">lung cancer investigation using expression profiling, a stem cell transplant investigation, biobanking is not an investigation, though it may be part of an investigation</IAO_0000011>
  <IAO_00000111 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">lung cancer investigation using expression profiling, a stem cell transplant investigation, biobanking is not an investigation, though it may be part of an investigation</IAO_0000011>
  <IAO_00000114 rdf:resource="http://purl.oblibrary.org/obo/OBI_00000122"/>
<IAO_00000115 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">a planned process that consists of parts: planning, study design execution, documenting, drawing a conclusion based on data, and which produce conclusions</IAO_00000115>
<IAO_00000117 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Joern Peters</IAO_00000117>
<IAO_00000119 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">BBI branch derived</IAO_00000119>
<IAO_00000232 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Could add specific objective specification</IAO_00000232>
<IAO_00000232 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Following OBI call November 2012,26th it was decided there was no need for adding 'achieves objective of drawing conclusion'; as existing relations were providing equivalent utility; this note closes the issue and validates the class definition to be part of the OBI core editor = PRS</IAO_00000232>
  
```

U:--- OBI-RDBBonesSubset.owl 30% L645 Git:master (nXML Valid)

Requirements



- Flexible User Interface For All Kinds Of Models (FUIFAKOM)
- Im- and Export of Data (and/or configuration files)
- Connectibility: I.e. the possibility to link different installations and make content comparable
- Fine grained user rights administration

Requirements

- Flexible User Interface For All Kinds Of Models (FUIFAKOM)
- Im- and Export of Data (and/or configuration files)
- Connectibility: I.e. the possibility to link different installations and make content comparable
- Fine grained user rights administration

Requirements



- Flexible User Interface For All Kinds Of Models (FUIFAKOM)
- Im- and Export of Data (and/or configuration files)
- Connectivity: I.e. the possibility to link different installations and make content comparable
- Fine grained user rights administration

Requirements



- Flexible User Interface For All Kinds Of Models (FUIFAKOM)
- Im- and Export of Data (and/or configuration files)
- Connectibility: I.e. the possibility to link different installations and make content comparable
- Fine grained user rights administration

Suitable frameworks for implementation

Options



UNI
FREIBURG

Options were:

- 1 Build an application from scratch (e.g. using java spring framework) or
- 2 Find a working and actively maintained implementation.

Suitable frameworks for implementation

Options



Options were:

- 1 Build an application from scratch (e.g. using java spring framework) or
- 2 Find a working and actively maintained implementation.

Suitable frameworks for implementation

Problems



- 1 Building from scratch was too ambitious**
- 2 There are (almost) no available free implementations. One oftentimes finds projects that served as proof-of-concept in a thesis but are abandoned afterwards, without reaching full functionality.**

Suitable frameworks for implementation

Problems



- 1 Building from scratch was too ambitious
- 2 There are (almost) no available free implementations. One oftentimes finds projects that served as proof-of-concept in a thesis but are abandoned afterwards, without reaching full functionality.

Abandoned Implementation projects

Linked Data frontend for SPARQL endpoints for Django

102 commits 2 branches 0 releases 1 contributor LGPL-3.0

Branch: master ▾ New pull request Create new file Upload files Find file Clone or download ▾

wikier updated general stuff Latest commit f081565 on 1 Sep 2014

apps/demo	fixed demo	7 years ago
doc/images	added a image describing pubby (inspired by pubby's one)	7 years ago
lib	updated general stuff	2 years ago
.gitignore	migrated ignored files from hg to git	2 years ago
COPYING.txt	license	7 years ago
ChangeLog.txt	fixed problem distributing templates on the tarball, issue #1	7 years ago
README.md	more info	2 years ago

README.md

This project is currently **not maintained**, so please use it under your own risk.

Djubby, a Linked Data frontend for SPARQL endpoints

Djubby is a Linked Data frontend for SPARQL endpoints for the Django Web framework. It's quite inspired by Richard Cyganiak's [Pubby](#), and with the exception of the HTML style, all the code has been written from scratch due to the many differences between languages (Java vs. Python) and the frameworks (JavaEE vs. Django).

HTML Browsers RDF Browsers SPARQL Clients





Abandoned Implementation projects

Screenshot of a GitHub repository page for "TheProjecter / django-rdf".

The repository has 1 issue, 0 pull requests, 0 projects, and no pulse activity.

The latest commit was made by [stebbi](#) on Jul 2008, with the commit hash `d37bf43`.

The commit message: "Removed the raw parameter from model save methods, as per recent Djan...".

The repository contains files: `examples`, `rdf`, `.svnignore`, `INSTALL`, `LICENSE`, and `README`.

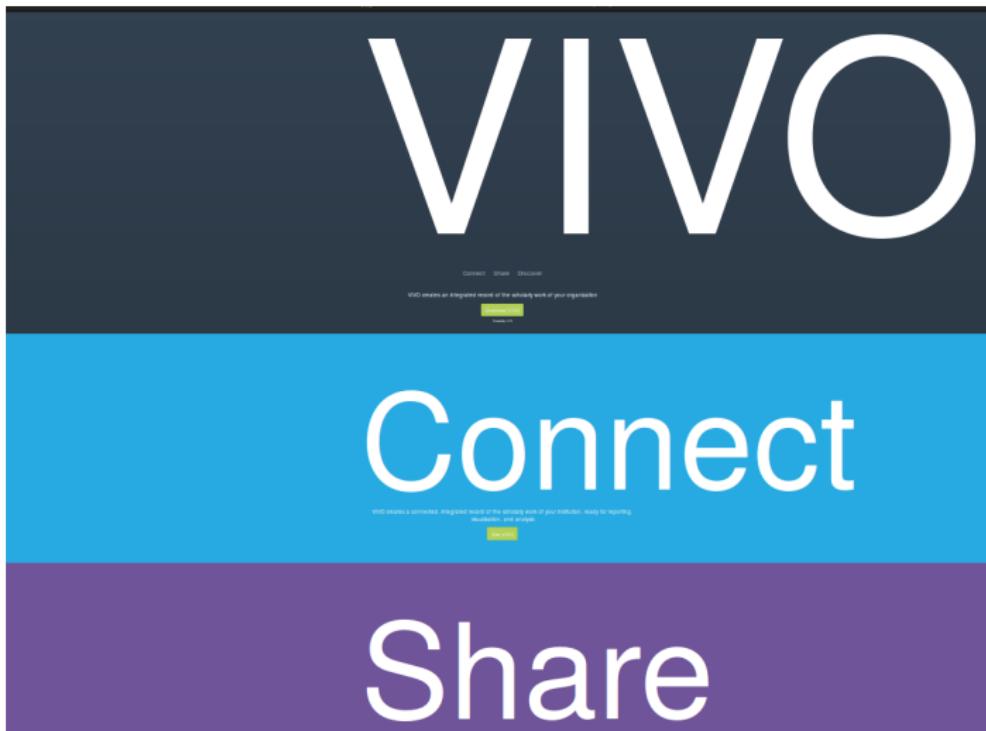
The `README` file content:

```
Nothing much here yet. Check out
http://code.google.com/p/django-rdf
instead...
```

Finally...



UNI
FREIBURG



The image shows a screenshot of the VIVO homepage. The top half has a dark grey header with the word "VIVO" in large white letters. Below the header is a light blue section containing the word "Connect" in large white letters. At the bottom is a purple section containing the word "Share" in large white letters. The center of the page features a yellow button with the text "Create my VIVO". The footer contains several navigation icons: back, forward, search, and others.

VIVO

Connect

Share

Create my VIVO

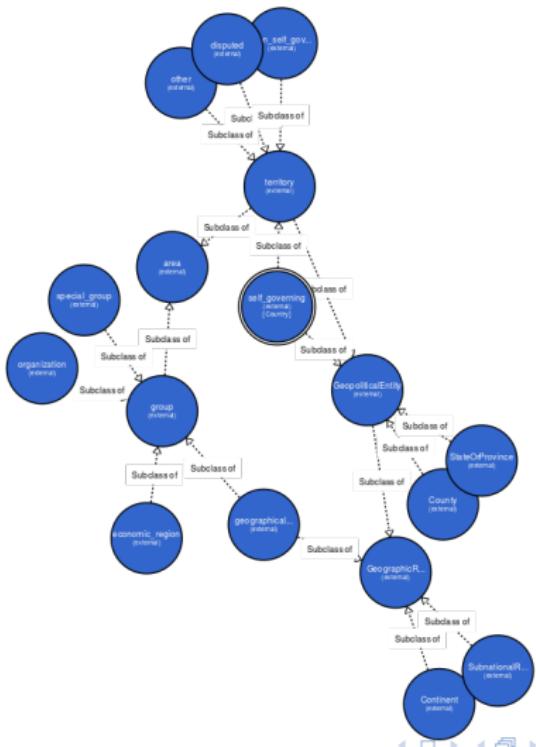
- Server-based web application
- Handles RDF-graphs
- Allows configuration using RDF, too

- Server-based web application
- Handles RDF-graphs
- Allows configuration using RDF, too

- Server-based web application
- Handles RDF-graphs
- Allows configuration using RDF, too

- The underlying network-graphs, provided by the ontologies, are represented by webpages (see problems later)
- Each object has a profile page, where the corresponding properties are listed

- The underlying network-graphs, provided by the ontologies, are represented by webpages (see problems later)
- Each object has a profile page, where the corresponding properties are listed



Resource URI: <http://aims.tao.org/aoe:geopolitical.owl#Afghanistan>



Afghanistan |  | Country 

Overview Identity File Other View All

in geographic grouping 

[Asia](#) Continent |  

[Economic Cooperation Organization](#) |  

[FAO](#) 2006 |  

[Food and Agriculture Organization of the United Nations](#) |  

[Food and Agriculture Organization of the United Nations](#) |  

[... more](#)

has border with 

[China](#) Country |  

[Iran \(Islamic Republic of\)](#) Country |  

[Pakistan](#) Country |  

Examples



Photo

Admin Panel | Edit this individual | Resource URI: <http://aims.fao.org/aos/geopolitical.owl#Asia> | Verbose property display is off | Turn on

Asia | | Continent | Transnational Region

Overview | Identity | File | Other | View All

Overview

has member country or territory

- [Afghanistan](#) Country |
- [Armenia](#) Country |
- [Azerbaijan](#) Country |
- [Bahrain](#) Country |
- [Bangladesh](#) Country |
- [Bhutan](#) Country |
- [Brunei Darussalam](#) Country |
- [Cambodia](#) Country |
- [China](#) Country |
- [Cyprus](#) Country |
- [Georgia](#) Country |

VIVO/Vitro - Data Input/Output

Data im- and export



UNI
FREIBURG

Vivo provides basic functionality to handle spreadsheet data. This functionality can be built upon for a more user friendly way to im- and export large amounts of existing data.



Ingest Menu > Convert CSV to RDF

comma separated tab separated

CSV file URL (e.g. "file:///")

Or upload a file from your computer:

No file selected.

This tool will automatically generate a mini ontology to represent the data in the CSV file. A property will be produced for each column in the spreadsheet, based on the text in the header for that column.

In what namespace should these properties be created?

Namespace in which to generate properties

Each row in the spreadsheet will produce a resource. Each of these resources will be a member of a class in the namespace selected above.

What should the local name of this class be? This is normally a word or two in "camel case" starting with an uppercase letter. (For example, if the spreadsheet represents a list of faculty members, you might enter "FacultyMember" on the next line.)

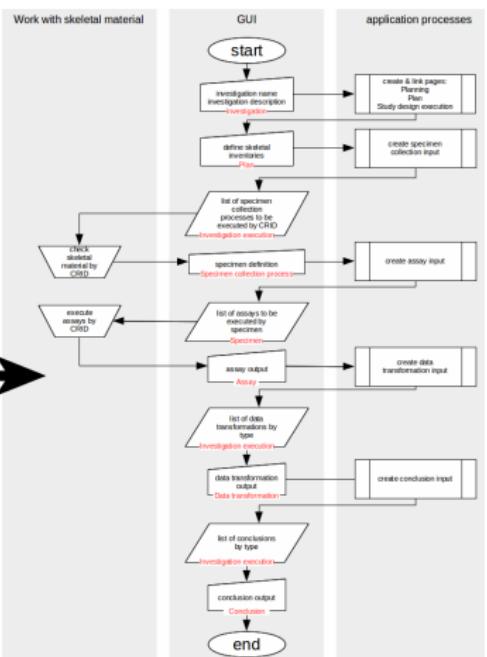
Class Local Name for Resources

From Ontology to workflow

Non-linearity vs. linearity



UNI
FREIBURG



From Ontology to workflow

Non-linearity vs. linearity



UNI
FREIBURG

While the ontology and the resulting web application exhibit a network structure with no distinct entry and exit points, a workflow for entering new data is usually a linear process (select data, do something in a consecutive manner)

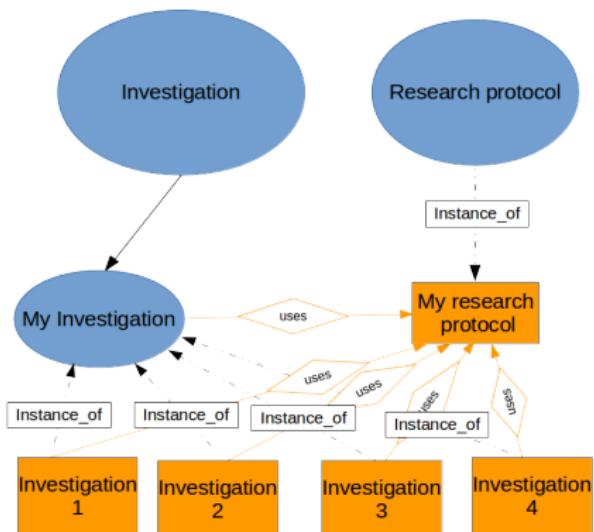
Automatic triple generation

When entering new data (creating instances) in VIVO, restrictions regarding <Class - Instance> relations are by default not respected:

Automatic triple generation



Example:



Automatic triple generation

Unfortunately, VIVO does not (yet) do this.

Solution

Fortunately, VIVO allows for easy (theoretically) modification and adaptability, without touching the actual VIVO code.