

A Bi-directional Semantic Framework for Bibliographic Metadata

John del Corral and M.Benno Blumenthal
International Research Institute for Climate
and Society

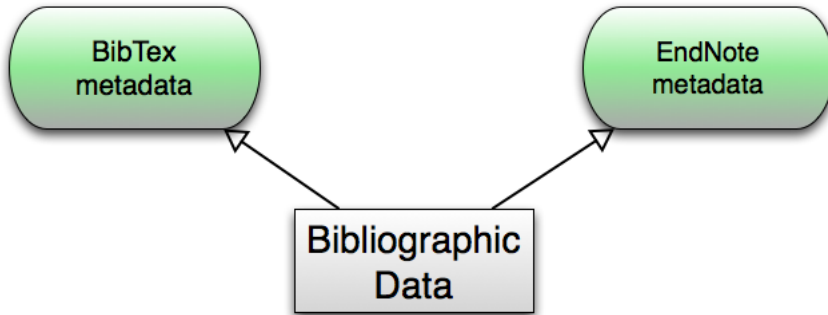
Columbia University

<http://iridl.ldeo.columbia.edu/>

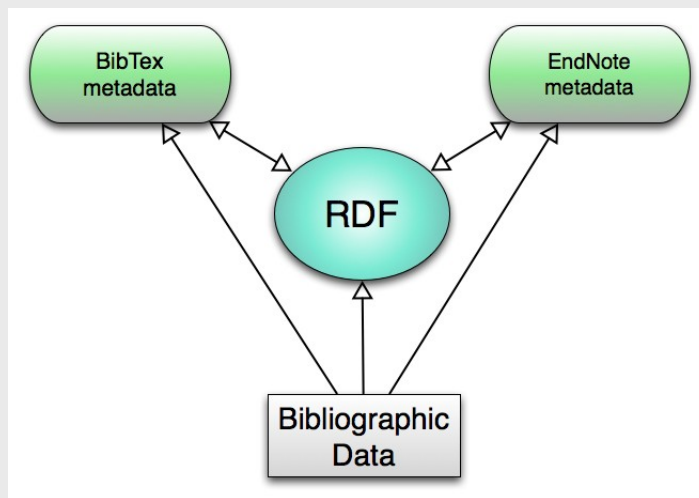
Motivation

- We have lots of data.
- People can not find what we have and what they want.
- We need better metadata.
- There are many standards for metadata.
- Which standard will meet our needs?
- Do we need to support multiple metadata standards?
- Which standard can be adapted to changing data types?

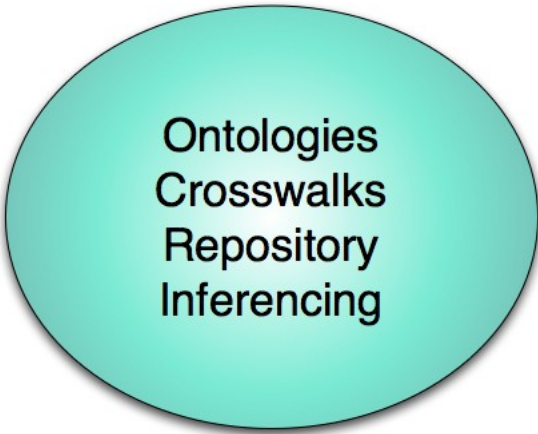
Metadata



Metadata + Resource Description Framework (RDF)



RDF

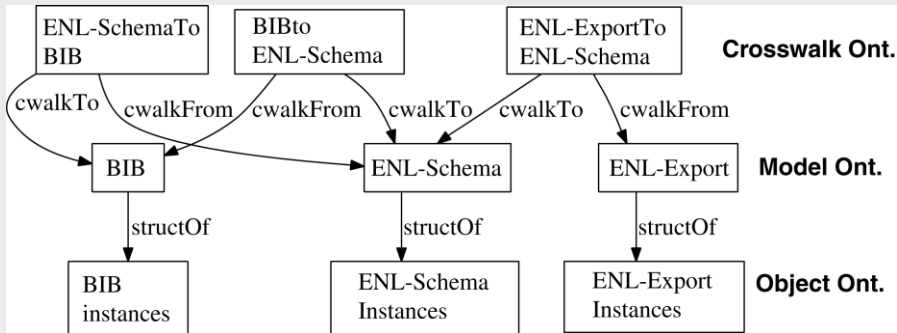


Ontologies
Crosswalks
Repository
Inferencing

Ontologies

- **Model Ontologies**
 - BibTeX
 - EndNote Export
 - EndNote Schema
- **Object Ontologies**
 - BibTeX Instances
 - EndNote Export Instances
 - EndNote Schema Instances
- **Crosswalk Ontologies**
 - EndNote Export To Endnote Schema
 - EndNote Schema To BibTeX
 - BibTeX To EndNote Schema

Ontology Mapping



Populating Object Ontologies

- BibTeX
 - BibTeX schema for Model
 - BibTeX-2-RDF for Instances
- EndNote Export
 - XML2OWL for Model
 - XML transform (.xsl) for Instances

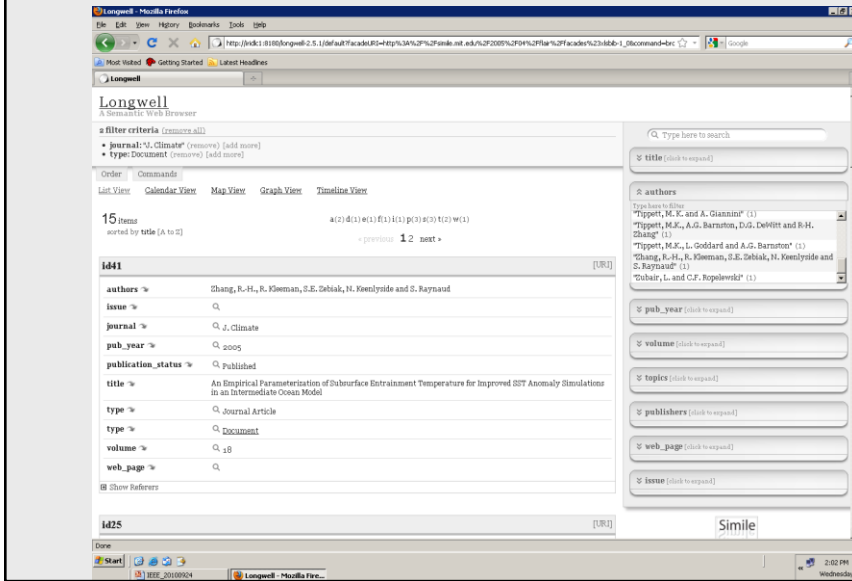
Crosswalks

- Web Ontology Language (OWL) Inference Rules
- Sesame RDF Query Language (SeRQL) CONSTRUCT Rules

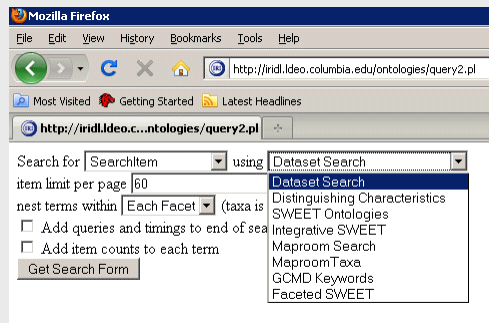
Inferencing

- Model Ontologies
- Crosswalk Ontology
- Triple Store Repository
- OWL Storage and Inference Layer (SAIL)
- Repository API To Iterate SeRQL CONSTRUCT Rules

MIT's Longwell Semantic Browser



Multiple Metadata Taxa



IRI Climate Data Library Faceted Search

INTERNATIONAL RESEARCH INSTITUTE

Search Results (24):

- DASILVA SMD94 anomalies outgoing longwave radiation**
anomalies outgoing longwave radiation from DASILVA SMD94: a five-volume atlas series depicting the seasonal and yearly variations of the surface marine atmosphere over the global oceans. Resolution: 1x1, Longitude: global, Latitude: global, Time: [Jan,1945,Dec,1993], monthly.
- DASILVA SMD94 climatology outgoing longwave radiation**
climatology outgoing longwave radiation from DASILVA SMD94: a five-volume atlas series depicting the seasonal and yearly variations of the surface marine atmosphere over the global oceans. Resolution: 1x1, Longitude: global, Latitude: global, Time: [Jan,Dec], monthly.
- DASILVA SMD94 halfbyhalf climatology outgoing longwave radiation**
halfbyhalf climatology outgoing longwave radiation from DASILVA SMD94: a five-volume atlas series depicting the seasonal and yearly variations of the surface marine atmosphere over the global oceans. Resolution: 0.5x0.5, Longitude: global, Latitude: global, Time: [Jan,Dec], monthly.
- Equatorial Pacific Pentad Outgoing Longwave Radiation Anomaly**
Pentad outgoing longwave radiation anomalies over the equatorial Pacific.
- Global Monthly Outgoing Longwave Radiation**
Monthly global observed and anomalous outgoing longwave radiation.

RDF Queries on data.gov.uk

INTERNATIONAL RESEARCH INSTITUTE

SPARQLing data.gov.uk: Edubase Data

10 October 2008, 11:16 am by Leigh Daniels in: Tutorials

Last week the Cabinet Office issued a call for Open Data Developers to sign-up to get a preview of the forthcoming UK Government public data website. The site includes a directory of existing datasets plus a growing number of datasets that have been converted to RDF and which will shortly be available as Linked Data. This data is being stored in the Talis Platform providing developers with access to SPARQL endpoints as a means to query the data, we'll also be including search and other access mechanisms at a later date.

In this series of postings I wanted to show some example SPARQL queries that can be used to access the data. If you're new to SPARQL then you might want to look at Lee Feigenbaum's SPARQL by Example tutorial, or my own short slide deck that covers all the basic syntax.

The first dataset I wanted to highlight is an extract of the Edubase dataset available from the Department of Children, Schools and Families. The conversion was carried out by the team at HP Labs and has been loaded into a Talis Platform store. The public facing SPARQL endpoint is available from: <http://services.data.gov.uk/edubase/0000/>

Here are some sample SPARQL queries you can use against the data.

Q1. Select the names of schools in the Administrative District of the City of London. # Ordering results by name of the school

```
prefix sch-ont: <http://education.data.gov.uk/det/school/>
SELECT ?name WHERE {
  ?school a sch-ont:School;
  sch-ont:establishmentName ?name;
  sch-ont:districtAdministrative <http://statistics.data.gov.uk/id/local-authority-district/0000/> ;
}
ORDER BY ?name
Results
```

Q2. Which schools in the BARNES area have a nursery?

```
prefix sch-ont: <http://education.data.gov.uk/det/school/>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?name WHERE {
  ?school a sch-ont:School;
  sch-ont:establishmentName ?name;
  sch-ont:districtAdministrative <http://statistics.data.gov.uk/id/local-authority-district/0000/> ;
  sch-ont:nurseryProvision "true"^^xsd:boolean
}
```

Ongoing Work

- Java-based Persistence Framework in Sesame2
- XSD2OWL for direct conversion of XML Schema to OWL ontologies
- OWL2 property chains

Notable References

- <http://data.gov.uk>
- <http://data.gov>
- <http://tw.rpi.edu>

Property Chain

ENL-Schema:author		Datatype property
mapstoPropertyChain =	BIB:author	Object property
	BIB:name	Datatype property