

Semantics System Overview

The Semantics Team:

Matthew B. Jones, Mark Schildhauer, Ben Leinfelder,
Xixi Luo, Deborah McGuinness, Lauren Walker, Chris
Jones, Rachel Volentine

February 6, 2015



Semantics: enabling improved discovery

Goal: Improve precision and recall of search

Annotating data with semantic types

Extending search systems to incorporate semantics

Display and edit semantic annotations in the web UI

Semantics Use Cases

Use Case 50	Automatic Annotation	Automatic assignment of measurement types via algorithm
Use Case 51	Manual Annotation	Manual assignment of measurement types by people
Use Case 52	Semantic Search	Improved precision and recall of data search via semantics
Use Case 53	Reject Annotation	Curation and rejection of annotations by data owners
Use Case 54	Annotation Conflicts	Resolution of conflicts from multiple annotations on one data
Use Case 55	Ontology Repository	A repository to store, version, and access ontologies
Use Case 56	Carbon Cycle Ontology	Knowledge Models for the carbon cycle and primary productivity

Semantics at-a-glance

New first-class objects in DataONE

Semantic Annotations

New Tools

New model for linking semantics to data

New KR models for Carbon Cycle and related areas

Ontology repository

Automated classifier to assign semantics to variables

Web UI for browsing and editing semantics



Measurement Annotation Display

Number Of
Records

3898583

Attribute Information

Variables

date

station

lat

lon

depth

temp

salinity

density

Name

density

Flag this tag as incorrect

Tags

Volumetric Density

Mass

Label

Definition

Ocean density

Storage Type

Measurement
Type

ratio

Measurement
Domain

Unit

kilogramPerCubicMeter

Type

real

1

3

+

Annotate

Number Of Records3898583

Attribute Information

Variables

date

station

lat

lon

depth

temp

salinity

density

Name

Tags

Label

Definition

Storage Type

Measurement Type

Measurement Domain

Missing Value Code

Accuracy Report

Accuracy Assessment

Coverage

density

Volumetric Density

Ocean density

ratio

Unitkilogram

Type

real

12

Sarah SmithApril 26, 2014 5:24 PM

Added a tag

Density

Definition: Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

from the SWEET ontology

Bill KertscherMay 25, 2014 5:24 PM

Added a tag

Volumetric Density

Definition: The ratio of the mass of any substance to the volume occupied by it (usually expressed in kilograms per cubic meter, but any other unit system may be used); the reciprocal of specific volume

from the SWEET ontology

Bill KertscherMay 25, 2014 5:24 PM

I chose this tag because consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit

Help others find this dataset by adding tags and comments

Add a comment...

Volume

Volume

Volumetric

Volumetric Heat

Volumetric Heat Capacity

Volume Fraction

Definition

Sed posuere consectetur est at lobortis. Aenean eu leo quam. Pellentesque ornare sem lacinia quam venenatis vestibulum.

People and Associated Pa

Data Set

Creators

Indiv

Search for Semantic Measurements

The screenshot displays the DataONE web interface. On the left, a map of the North Pacific Ocean is shown with a heatmap overlay representing data density. The map includes labels for the Bering Sea, Sea of Okhotsk, and North Pacific Ocean. A search bar at the top center contains the text "Search for data". To the right of the search bar, the DataONE logo and navigation links "DATA" and "ABOUT" are visible. Below the search bar, a panel titled "Datasets 1 to 25 of 267" displays a list of datasets. The datasets are sorted by "Most recent". The first dataset listed is "Historical CTD data from the Gulf of Alaska, 1970-2010" by Thomas J. (Tom) Weingartner, Thomas Royer, and Russell Hopcroft. The second dataset is "Relationship of Community Characteristics to Harvest Reporting: Comparative study of household surveys and harvest tickets in Alaska" by Jennifer I. Schmidt and Bonanza Creek LTER. The third dataset is "seabird and mammal surveys in Kenai Fjords National Park, Alaska: 2006-2008, PWS: 2012" by Mayumi Arimitsu. The fourth dataset is "Additional sites added to the Delta Junction fire-chronosequence soil database: carbon, nitrogen, bulk density, and other properties" by Kristen L Manies, Jennifer W. Harden, and Bonanza Creek LTER. The fifth dataset is "The impact of permafrost thaw on ecosystem carbon balance: Seasonal water table depth from 2012-2013" by Edward A.G. Schuur, Christina Schadel, and Bonanza Creek LTER. The sixth dataset is "The" by Edward A.G. Schuur, Caitlin Elizabeth Hicks-Brien, and Bonanza Creek LTER. On the left side of the map, a "Filter" panel is open, showing a search bar with the text "Anything". Below the search bar, a "Data attribute (density)" panel is open, showing a list of attributes: Temperature, Time, Type, Count, Length, and Proportion. A "Definition" panel is also open, showing the text "Sed posuere consectetur est at lobortis. Aenean eu leo quam. Pellentesque ornare sem lacinia quam venenatis vestibulum." At the bottom of the map, a "Data covers" panel is open, showing a date range from 1900 to 2014 and a "Data covers" checkbox.

Search for data

DataONE DATA ABOUT

Datasets 1 to 25 of 267

1 2 3 ... 11 Next

Sort by Most recent

Thomas J. (Tom) Weingartner, Thomas Royer, and Russell Hopcroft. 2014. **Historical CTD data from the Gulf of Alaska, 1970-2010.** (doi:10.5063/F1JM27JF)

Jennifer I. Schmidt, and Bonanza Creek LTER. 2014. **Relationship of Community Characteristics to Harvest Reporting: Comparative study of household surveys and harvest tickets in Alaska.** knb-lter-bnz.578.2

Mayumi Arimitsu. 2014. **seabird and mammal surveys in Kenai Fjords National Park, Alaska: 2006-2008, PWS: 2012.** df35b.219.11

Kristen L Manies, Jennifer W. Harden, and Bonanza Creek LTER. 2013. **Additional sites added to the Delta Junction fire-chronosequence soil database: carbon, nitrogen, bulk density, and other properties.** knb-lter-bnz.540.3

Edward A.G. Schuur, Christina Schadel, and Bonanza Creek LTER. 2013. **The impact of permafrost thaw on ecosystem carbon balance: Seasonal water table depth from 2012-2013.** knb-lter-bnz.554.3

Edward A.G. Schuur, Caitlin Elizabeth Hicks-Brien, and Bonanza Creek LTER. 2013. **The**

Filter

Anything

Data attribute (density)

Temperature

Time

Type

Count

Length

Proportion

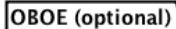
Definition

Sed posuere consectetur est at lobortis. Aenean eu leo quam. Pellentesque ornare sem lacinia quam venenatis vestibulum.

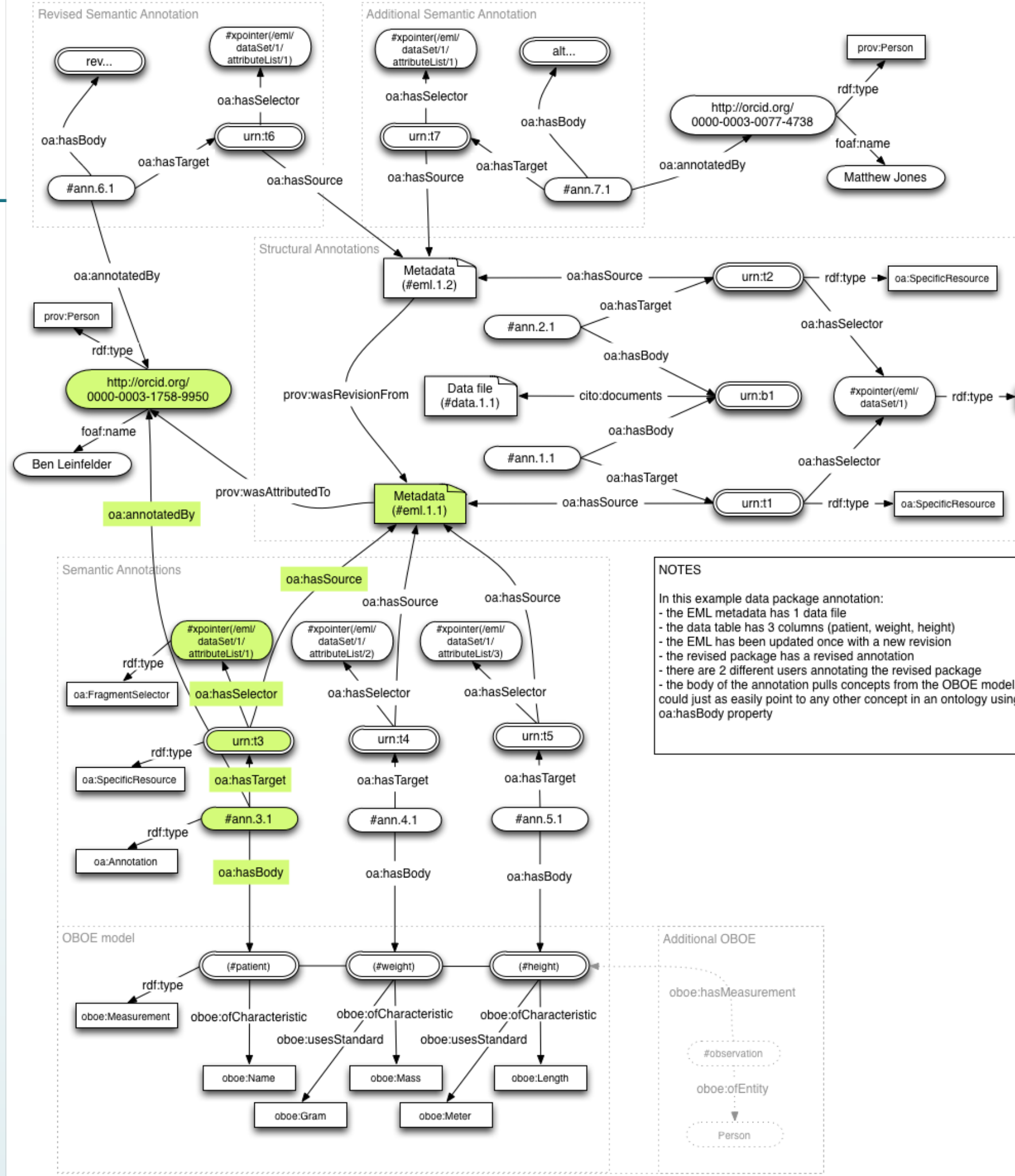
1900 2014

Data covers

tagging



Semantic Annotation Example



Carbon Cycle Ontology



Semantics status (as of Feb 1, 2015)

Design documentation

First draft completed

Annotation Model

Drafted, getting feedback

Indexing and Search

In progress

Web display

Mockups done, prototype development started

Carbon Cycle Model

Initial draft, in revision

Ontology repository

Using BioPortal