OOI Semantic Prototype

http://www.oceanobservatories.org/spaces/display/CIDev/Semantic+Framework+Integration

Luis Bermudez (SURA)
bermudez@sura.org
2009.11.16

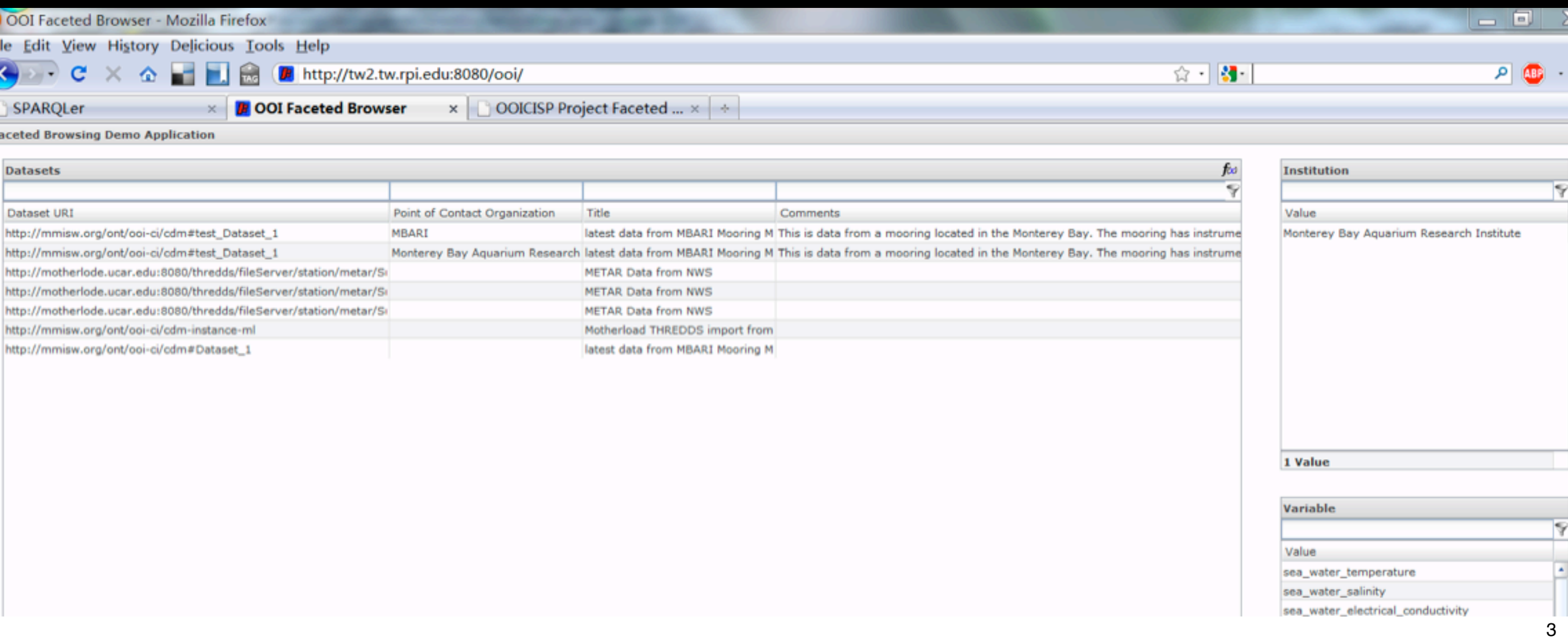
Scenario

From "ORION Cyberinfrastructure Concept of Operations - Version 0r08 2006.05.11", Section B: Getting and Using Products - (page 6)

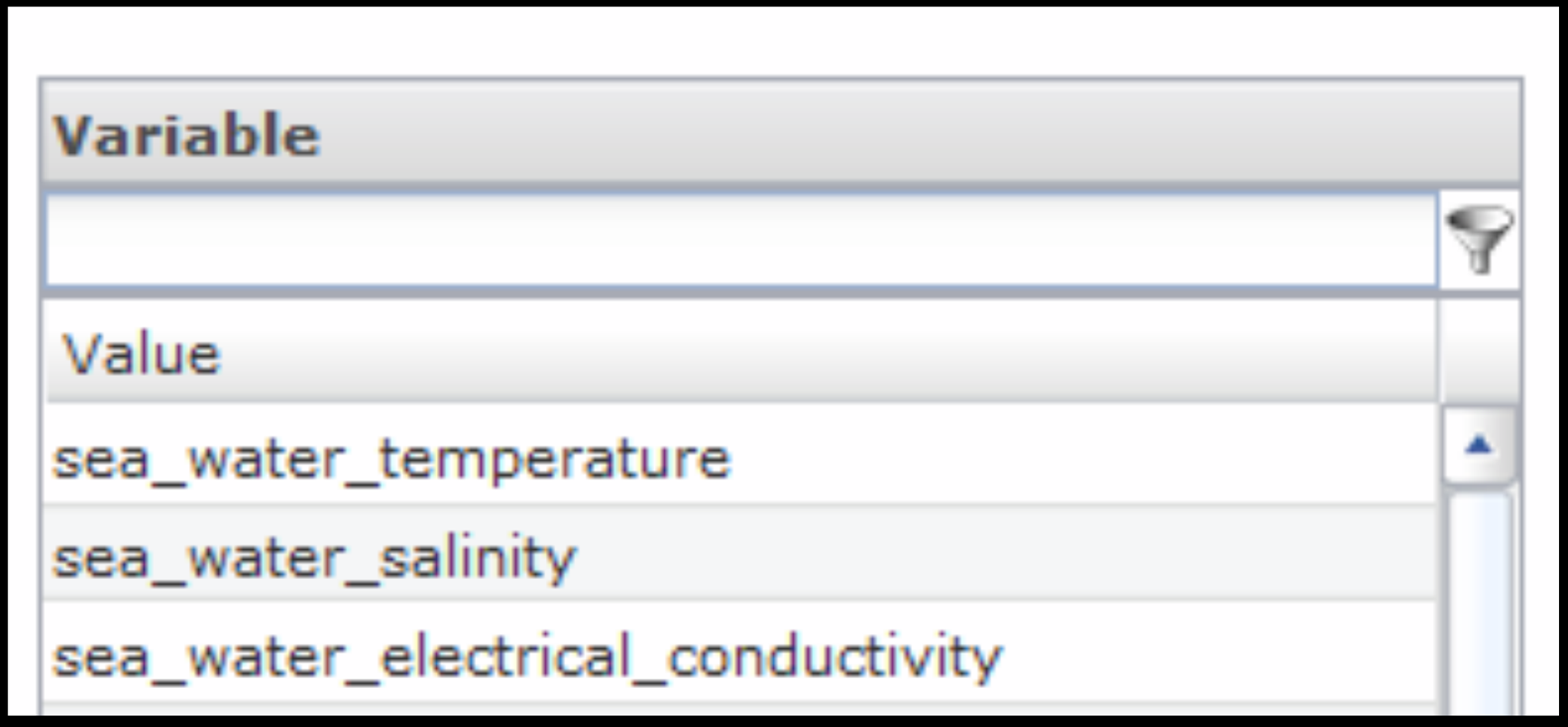
... (POIM refers to the model that Dr. Chu is building and using. It assimilates data to run the model.)

Fortunately, the ORION systems all support a straightforward "publish and subscribe" model for distributing data. A data user can subscribe to a data stream, asking to receive any new data as soon as they is measured by an instrument. In this case, POIM has subscribed to all of the variables it needs, and executes whenever a new value is received for any of them. If data arrive while POIM is running, it will cache them, finish

Categorization of distributed heterogenous data



Categorization or Facet



Sources examples:

Top level IOOS Regional Modeling catalog:

Description: Top level catalog of Regional Association catalogos (AOOS, NANOOS, etc..). It includes not only recent but historic numerical data.

Link: http://coast-enviro.er.usgs.gov/thredds/ioos_catalog_top.html (xml)

NDBC Observed Data

Description: Data from different projects including Ocean Sites, TAO, NDBC buoys etc..

Link: http://dods.ndbc.noaa.gov/thredds/catalog.html (xml)

http://www.oceanobservatories.org/spaces/display/CIDev/Data+Sources

Heterogeneity

Different Long Names

time: Array of 64 bit Reals [time = 0..24]

time:

```
long_name: "Valid Time"
units: "hour since 200
NAVO_code: 13
time_origin: "2000-01-
```

http://edac-dap.northerngulfinstitute.org/thredds/dodsC/ncom/region7/ncom_glb_reg7_2009110300.nc.html

time: Array of 32 bit Integers [time = 0..22015]

time:

```
long_name: "Epoch Time"
short_name: "time"
standard_name: "time"
units: "seconds since 1970-01-01 00:00
```

http://dods.ndbc.noaa.gov/thredds/dodsC/data/ocean/15319/15319o9999.nc.html

No Standard Name attribute

No NAVO code

Incomplete Metadata

```
institution: "NOAA National Data Buoy Center and
Participators in Data Assembly Center"
url: "http://dods.ndbc.noaa.gov"
quality: "Automated QC checks with periodic manual QC"
conventions: "COARDS"
station: "15319"
comment: "Sanha - WANE 25947"
location: "5.58 S 11.83 E "
Unlimited_Dimension: time
```

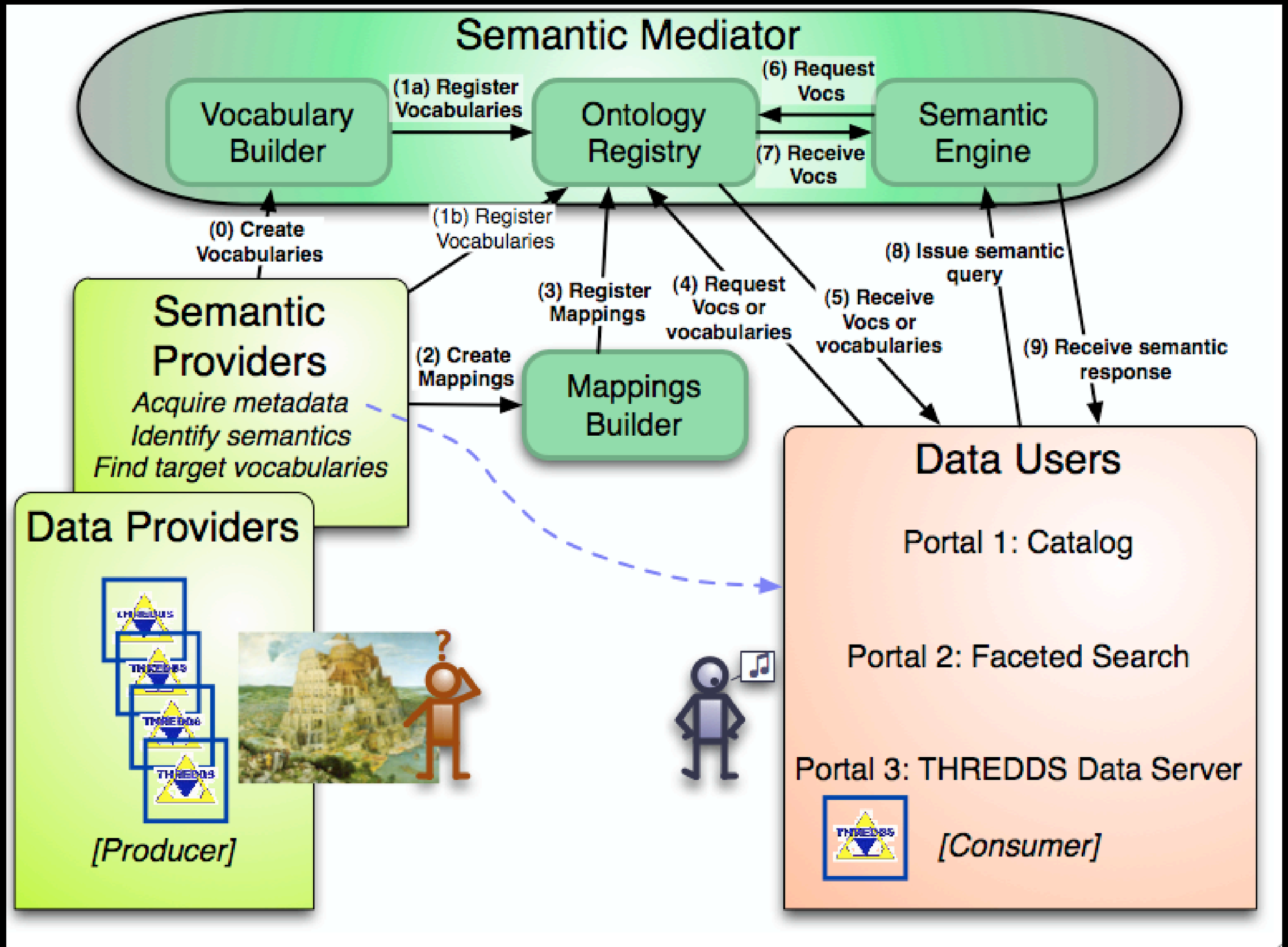
Contact
Information ?

Different CF Metadata Conventions

```
operational_status: "development"
model_type: "x-curvilinear lon, y-
input_data_source: "FNMOC NOGAPS,
Conventions: "NAVO_netcdf_v1.0"
message: "UNCLASSIFIED"
reference: "https://www.navo.navy.
```

Global Attributes:

```
institution: "NOAA Nati
Data Assembly Center"
url: "http://dods.ndbc.
quality: "Automated QC
conventions: "COARDS"
station: "15319"
```


Semantic Mediator
Vocabulary Builder
(1a) Register Vocabularies
Ontology Registry
(6) Request Vocs
(7) Receive Vocs
Semantic Engine
(0) Create Vocabularies
(1b) Register Vocabularies
(3) Register Mappings
(4) Request Vocs or vocabularies
(5) Receive Vocs or vocabularies
(8) Issue semantic query
(9) Receive semantic response
Semantic Providers
Acquire metadata
Identify semantics
Find target vocabularies
(2) Create Mappings
Mappings Builder
Data Users
Portal 1: Catalog
Portal 2: Faceted Search
Portal 3: THREDDS Data Server
[Consumer]
Data Providers
[Producer]

Leverage MMI

The MMI Ontology Registry and Repository

Register

Marine Metadata Interoperability

Ontology Registry and Repository alpha

bermud

Review and Register | Cancel

(creating new ontology) -

▶ Metadata details

▼ Contents

Synopsis of ontology contents:

Please specify your ontology file.

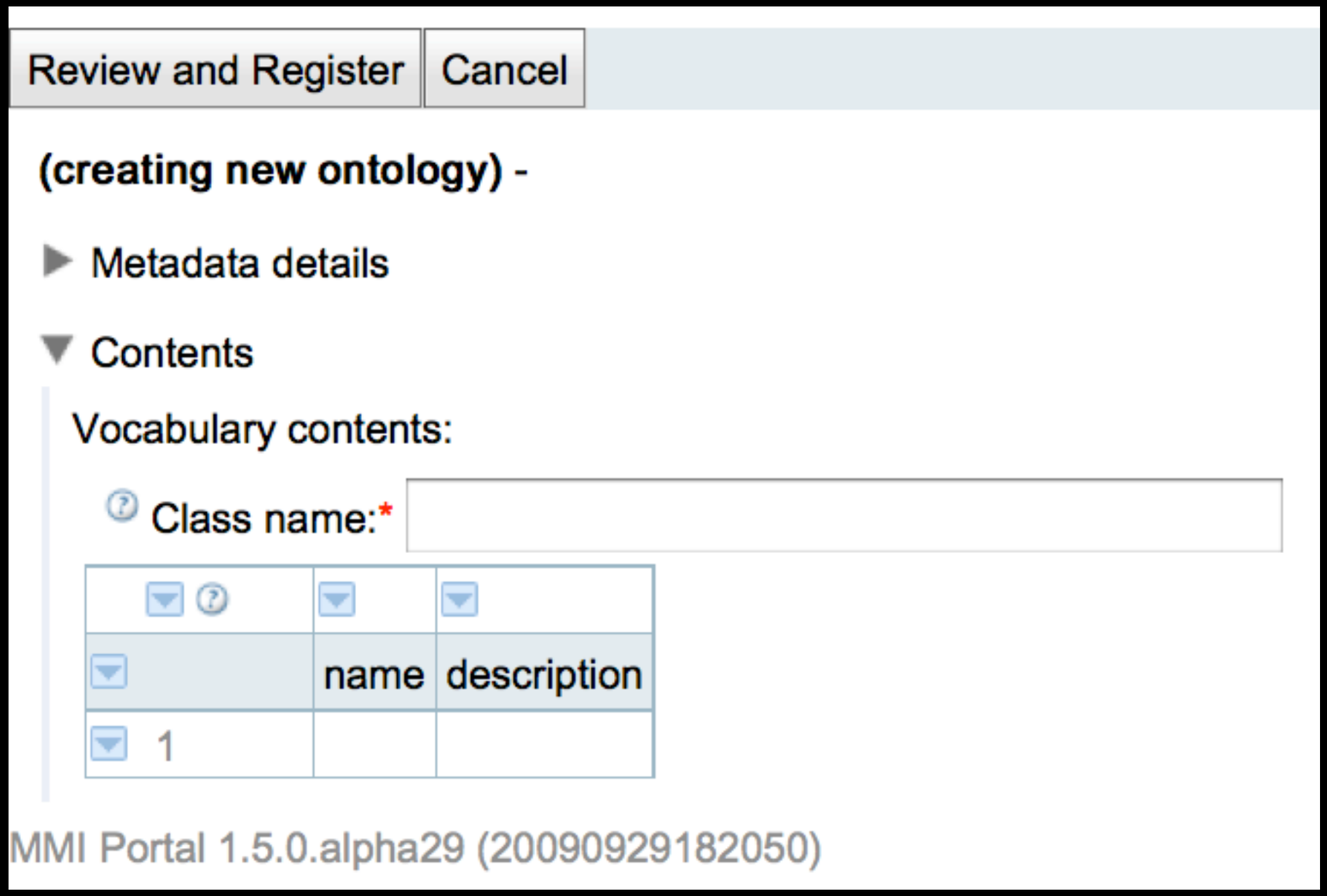
◉ Local file: Browse...

☐ Preserve original base namespace

Load ontology | Details

MMI Portal 1.5.0.alpha29 (20090929182050)

Create ontologies from scratch



Metadata details

General | **Usage/License/Permissions** | **Original source**

Resource type:	qualityFlag (http://mmisw.org/ont/mmi/theme/qualityFlag)
Full title:	Argo QA/QC Flags
Content creator:	
Ontology creator:	Stephanie Watson
Brief description:	QA/QC flags used in Argo
Keywords:	quality, QA, QC, Argo
Link to original vocabulary:	
Link to documentation:	http://argo.jcommops.org/FTPRoot/Argo/Doc/argo-quality-control-ma
Authority abbreviation:	argo
Contributor(s):	Annie Wong, Robert Keeley, Thierry Carval, and the Argo Data Management Team.

Update and Revisions

Available versions for http://mmisw.org/ont/argo/qualityFlag

URI	Name	Author	Version	Submitter
http://mmisw.org/ont/argo/20081116T040146/qualityFlag	Argo QA/QC Flags	Stephanie Watson	20081116T040146	graybeal
http://mmisw.org/ont/argo/20081113T203523/qualityFlag	Argo QA/QC Flags	Stephanie Watson	20081113T203523	swatson

Close

Browse

Marine Metadata Interoperability

Ontology Registry and Repository alpha

Create vocabulary | Upload

- Submitter
  - -All-
  - aisenor
  - amilan
  - bermud
  - carueda
  - carueda3
  - caruedagm
  - cchandler
  - cdip
  - dallison
  - dhr.sfsu@gmail.com
  - graybeal
  - haag
  - haines
  - jfredericks
  - kwalz

Name	Author
AGU Index Terms	AGU
ARGO Instruments	ARGO
ARGO Parameters	ARGO
Argo QA/QC Flags	Stephanie Watson
Authority Vocabulary	MMI
CDIP Term Vocabulary	CDIP
CeNCOOS water monitoring subset and extension of CF parameter vocabulary	Dale Robinson
CF NetCDF Standard Names	CF Metadata
Climate and Forecast (CF) standard names parameter vocabulary	Luis Bermudez
Demo in the q2o break outT	q2o
DRDC and CF Mapping	Anthony W. Isenor
DRDC Atlantic NADAS Parameter Codes	Anthony W. Isenor
EPIC Key File (Units)	Pacific Marine Environmental Laborat (NOAA)
Extended Contintental Shelf (ECS) Device Vocabulary	Amilan
Integrated Global Ocean Services System (IGOSS) QA/QC Flags	Stephanie Watson
IOOS Vocabulary Version 1	Stephanie Watson
IRD QA/QC Flags	Stephanie Watson

Use - Get the RDF

http://mmisw.org/ont?form=rdf&uri=http://mmisw.org/ont/argo/qualityFlag


```
    <omv:description>QA/QC flags used in Argo</omv:description>
    <omv:version>Latest terms per null</omv:version>
    <omvmmi:origMaintainerCode>argo</omvmmi:origMaintainerCode>
    <omv:keywords>quality, QA, QC, Argo</omv:keywords>
    <omv:hasContributor>Annie Wong, Robert Keeley, Thierry Carval, and the Argo Data M
    <omv:acronym>qualityFlag</omv:acronym>
    <omv:uri>http://mmisw.org/ont/argo/qualityFlag</omv:uri>
    <omv:documentation>http://argo.jcommops.org/FTPRoot/Argo/Doc/argo-quality-control-
    <dc:description>QA/QC flags used in Argo</dc:description>
    <omvmmi:shortNameUri>http://mmisw.org/ont/mmi/theme/qualityFlag</omvmmi:shortNameU
    <dc:title>Argo QA/QC flags</dc:title>
    <dc:contributor>Annie Wong, Robert Keeley, Thierry Carval, and the Argo Data Manag
    <omvmmi:creditRequired>not specified</omvmmi:creditRequired>
</owl:Ontology>
<owl:Class rdf:about="http://mmisw.org/ont/argo/qualityFlag/Qualityflag">
    <rdfs:label>qualityflag</rdfs:label>
</owl:Class>
<owl:DatatypeProperty rdf:about="http://mmisw.org/ont/argo/qualityFlag/code">
    <rdfs:domain rdf:resource="http://mmisw.org/ont/argo/qualityFlag/Qualityflag"/>
    <rdfs:label>Code</rdfs:label>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:about="http://mmisw.org/ont/argo/qualityFlag/description">
    <rdfs:domain rdf:resource="http://mmisw.org/ont/argo/qualityFlag/Qualityflag"/>
    <rdfs:label>Description</rdfs:label>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:about="http://mmisw.org/ont/argo/qualityFlag/source_notes">
    <rdfs:domain rdf:resource="http://mmisw.org/ont/argo/qualityFlag/Qualityflag"/>
    <rdfs:label>Source Notes</rdfs:label>
</owl:DatatypeProperty>
<Qualityflag rdf:about="http://mmisw.org/ont/argo/qualityFlag/_7">
    <source_notes>Argo from from IODE Summary Spreadsheet QualityFlags.xls</source_not
    <description>Not used</description>
    <code>7</code>
```

SPARQL Queries

```
SELECT ?property ?value
WHERE
{<urn:ogc:def:datatype:iso-19107::gm_polygon>
?property ?value. }
```

Marine Metadata Interoperability

Ontology Registry and Repository

alpha

Your SPARQL query:

Submit

SPARQL (Simp
query language
information her
Use this form to
Registry and R
Both GET and

CONSTRUCT examples	
`CONSTRUCT { ?s ?p ?o } WHERE {?s ?p ?o. } LIMIT 20`	`SELECT ?s ?p`
Submit	Submit

`PREFIX dc: <http://purl.org/dc/elements/1.1/>`
`CONSTRUCT { ?s dc:creator ?creator }`
`WHERE { ?s dc:creator ?creator. }`
	`PREFIX dc: <ht`
	`SELECT ?s ?c`
	`WHERE { ?s dc:`

Submit	Submit

`PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>`
`PREFIX owl: <http://www.w3.org/2002/07/owl#>`
`CONSTRUCT { ?class rdf:type owl:Class. }`
`WHERE { ?class rdf:type owl:Class. } LIMIT 20`
	`PREFIX rdf: <h`
	`PREFIX owl: <h`
	`SELECT ?class`
	`WHERE { ?class`

http://mmisw.org/ont/sparql.html

SPARQL Result


```xml
<?xml version="1.0"?>
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:j.0="http://mmisw.org/ont/ogc/def/">
    <j.0:Ogcdef rdf:about="urn:ogc:def:datatype:iso-19107::gm_polygon">
        <j.0:uri>urn:ogc:def:dataType:ISO-19107::GM_Polygon</j.0:uri>
        <rdfs:label>urn:ogc:def:dataType:ISO-19107::GM_Polygon</rdfs:label>
        <j.0:description>A Polygon is a special surface that is defined by a
            single surface patch</j.0:description>
    </j.0:Ogcdef>
</rdf:RDF>
```

Handling of URNs

http://mmisw.org/portal/#http://mmisw.org/ont/ogc/def

	uri	description
1	urn:ogc:def:crs:OGC:1.3:CRS83	NAD83 longitude-latitude B.4 in OGC 06-042
2	urn:ogc:def:dataType:ISO-19107::GM_Aggregate	A (heterogeneous) geometry collection that includes one or more geometry members
3	urn:ogc:def:ebRIM-ClassificationScheme:UNSD::GlobalRegions	Standard country or area codes and geographical regions for statistical use (UNSD)
4	urn:ogc:def:crs:OGC:1.3:AUTO42003:99:8888	Auto orthographic B.9 in OGC 06-042
5	urn:ogc:def:ebRIM-RegistryPackage:OGC::Root	Root package containing all extension packages
6	urn:ogc:def:ebRIM-RegistryPackage:OGC::Basic	Basic extension package
		Describes the service endpoints and

Easy Mapping

Working ontologies: Add...

A: http://mmisw.org/ont/igoss/qualityFlag/ -- *Integrated Global Ocean Services System (IGOSS) QA/QC*
B: http://mmisw.org/ont/argo/qualityFlag/ -- *Argo QA/QC Flags*

Search the following ontologies: A B

Search for: REGEX

Select: All None Selected: 1 out of 14 element(s)

- A:_7
- A:_5
- A:_4
  - http://mmisw.org/ont/igoss/qualityFlag/_4
  - **label**: 4
  - **code**: 4
  - **description**: The element appears erroneous
  - **source_notes**: Integrated Global Ocean Services System from from IODE Summary Spreadsheet

= ≈ < > ~

Config

Search the following ontologies: A B

Search for: REGEX

Select: All None Selected: 1 out of 14 element(s)

- B:_0
- B:_6
- B:_5
- B:_3
  - http://mmisw.org/ont/argo/qualityFlag/_3
  - **label**: 3
  - **code**: 3
  - **description**: Probably bad data (spike-gradient- ? if other tests passed)
  - **source_notes**: Argo from IODE Summary

Select: All None Mappings:

A:_4. ≈ B:_3.

= Exact match
≈ Close match
< Narrower than
> Broader than
~ Related match

Reducing risk within MMI

MMI collaborates with:

- Open Ontology Repository Initiative
- Bio Portal
- NeOn

Link to registry: http://mmisw.org

Projects that are using the semantic mediator

- Semantic Framework Integration (OOI)
- International Coastal Atlas Network (ICAN)
- OGC Ocean Science Interoperability Experiment
- Oceans Innovation Demo 2008
- Q2O (QARTOD 2 OGC)

Risk Reduction

2.13 2226 Comprehensive Semantic Framework

Risk description: If a comprehensive semantic framework cannot be devised then data interoperability between communities using different vocabularies will be limited. The ability of a broad community of ocean scientist to utilize many types of data requires the ability to mediate diverse data types. A powerful way to provide this is to allow actors utilizing different vocabualries to have a common semantic understating.

- **Starting status**
  - Likelihood =3 = Possible
  - Consequence = 4 = (Technical) Marginally acceptable, barely able to perform needed science. Impact of 5-10%
  - Rating = medium - some disruption is likely to occur.
- **Ending Status Goal:**
  - Likelihood =1 = Remote
  - Consequence =3 = (Technical) Acceptable, with significant reduction capability. Impact of 1-5%
  - Rating = low - minimum impact

Risk Reduction Tasks

- **Tasks:**
  - Realized the framework by designing / selecting and testing possible components
    - Define components:
      - Semantic Registry
      - Semantic Enabled User interface - faceted search
      - Metadata enhancer - semantic wiki
      - Data2knowledge converter
      - Data / Metadata Harvester
      - Service Registry

Working Group

- 4 organizations
- more than 10 participants
- Different interests

- John Graybeal (lead), UCSD
- Luis Bermudez (co-lead), SURA
- Peter Fox (principal), RPI
- Ilya Zaslavsky (principal), UCSD
- Carlos Rueda (developer), MBARI
- Matthew Rodrigez (developer), UCSD
- Michael Meisinger (architect), UCSD
- Roger Unwin (developer), UCSD
- Stephan Zednik (developer), RPI
- Zhenning Shangguan (developer), RPI
- Patrick West (developer), RPI
- Baoije Iowa (developer), RPI

Logistics goals:

- Share and agree on same goal, final picture
- Iterative process of What, how, who, when

Process Progress

- Identified Components - **done** !
- Select technologies for each component - **done**!
- Fast prototyping and creating test artifacts - **in progress**
  - Revise - Refine - Revise - Refine.....

Components

- Semantic Registry and Repository MBARI
- Harvester UCSD
- Faceted Browser RPI
- Metadata Registry UCSD
- Semantic Mediawiki UCSD

Supporting Tools

- Mailing List - ideally all communication should take place in **the** list.
- Wiki - general documentation, meetings, **meeting TODOS**, public dissemination, and conclusions.
- Weekly meetings - At least one per week. Summary updates, discuss issues, and all-on-the-same-page checking.
- One or more ad-hoc meetings during the week: Issues, refine ideas and next steps.

Example Wiki Topics

- CF Compliance
- Data Sources
- Facets for the UI
- Interaction between system components
- Main Components of Semantic Prototype
- Meetings for the Semantic Prototype
- Ontologies
- OOI CI Semantic Risks
- Prototype for Semantic Framework (deprecated)
- Risk Info (Summarized for Semantic Framework Team)
- Scenario from Concept of Observation
- Work Descriptions for Semantic Prototype

Supporting Tools

- JIRA (task management) will answer:
  - Have we done what we promise ?
  - Reporting
    - Main component level
    - SOW tasks
    - Bi-weekly reports

JIRA User: Luis Bermudez History | Filters | Profile | Log Out

HOME BROWSE PROJECT FIND ISSUES CREATE NEW ISSUE ADMINISTRATION

Issue Details (XML | Word | Printable)

Key:	CIDEVDM-41
Type:	Sub-task
Status:	Resolved
Resolution:	Fixed
Priority:	Major
Assignee:	Carlos Rueda
Reporter:	Luis Bermudez
Votes:	0
Watchers:	0

Available Workflow Actions

- Verify Issue
- Reopen Issue

Operations

- Assign this issue (to me)
- Attach file to this issue
- Attach screenshot to this issue
- Clone this issue
- Comment on this issue
- Delete this issue
- Edit this issue
- Move this sub-task

CI Development—DM
↳ Demonstate ability to register and retrieve ontologies (CIDEVDM-8)
↳ **Create screen cast capturing the browsing and downloading process of ontologies at the MMI Registry**

Created: 07/Nov/09 02:42 AM Updated: Saturday 09:15 PM

Return to search

Issue **8** of **29** issue(s)
<< Previous | CIDEVDM-41 | Next >>

Component/s:	Semantic Framework Integration Prototype
Affects Version/s:	R1I1
Fix Version/s:	None

Time Tracking:		
	Original Estimate:	6 hours
	Remaining Estimate:	4 hours
	Time Spent:	2 hours

All | Comments | Work Log | Change History | FishEye Sort Order:

Carlos Rueda - 14/Nov/09 08:58 PM [Permlink]

Time Worked: 2 hours

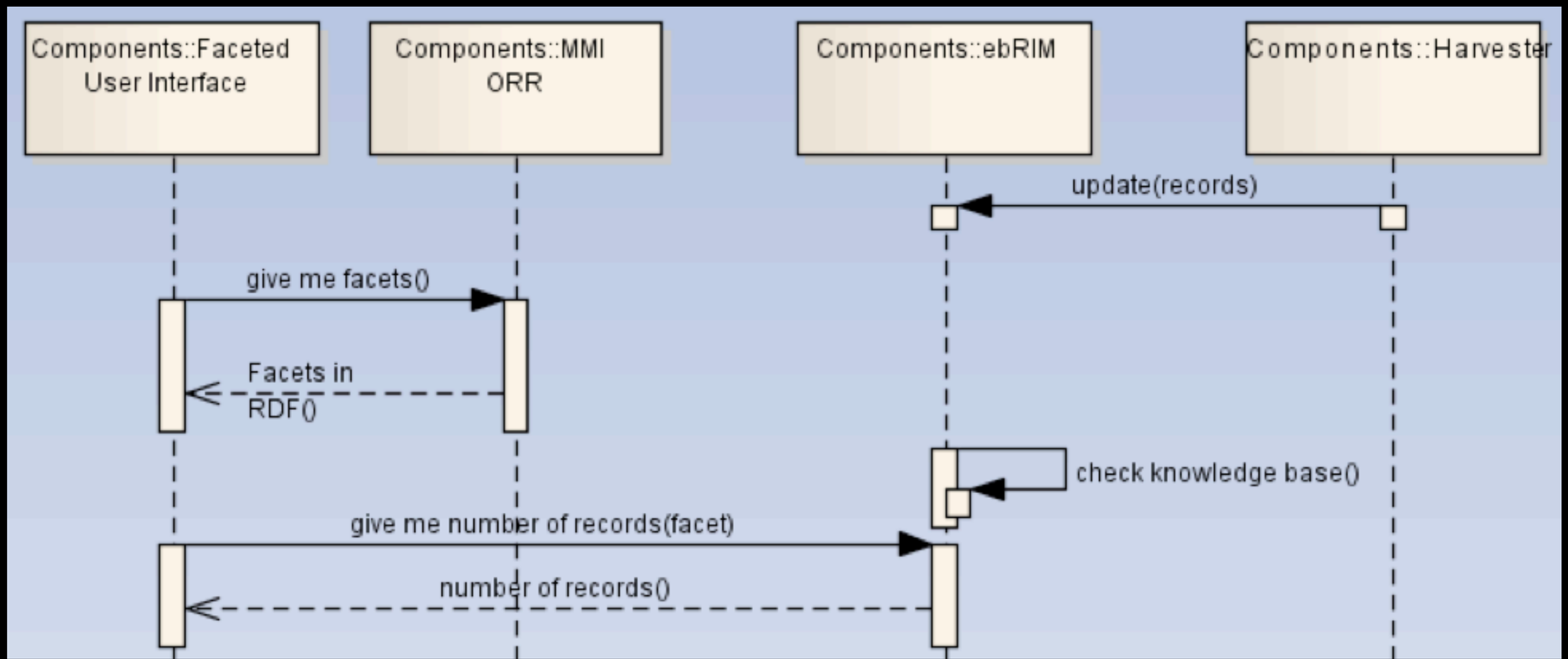
MMI ORR manual page updated: http://marinemetadata.org/mmiorrusrman/mmiorrnewusracct/mmiorrsignin/mmiorrnewvocab/mmiorrviewvocab
(link to the screencast included there).

http://ci.oceanobservatories.org/tasks/browse/CIDEVDM-41?page=com.atlassian.jira.plugin.system.issuetabpanels%3Aworklog-tabpanel

Supporting Tools

- UML - communicating software artifacts
  - Use cases for demonstration of functionality
  - Sequence diagrams
  - ...
- Code Repository (SVN, GIT)
  - share software artifacts : UML, ontologies, software code, ...

Example UML



http://www.oceanobservatories.org/spaces/display/CIDev/Interaction+between+system+components

Questions ?

Suggestions ?