Enabling Semantics in the OPeNDAP

Luis Bermudez

OPeNDAP Developers meeting - Feb 2007

Outline

- Knowledge Construction and RDF
- URIfying controlled vocabularies
- Practical use : OPeNDAP as OGC SOS

Why do we need semantics ?

We should be
able to identify, define and
have thesauri (ontologies)
for:

- Earth Realm
- Physical Phenomena
- Physical Processes
- Physical Properties
- Physical Substances
- Data Centers
- Human Activities
- Material Things
- Sensors
- Units
- Quality Flags
- etc.

Semantic Web

The internet is an electronic Library of Congress without a cataloguing system.

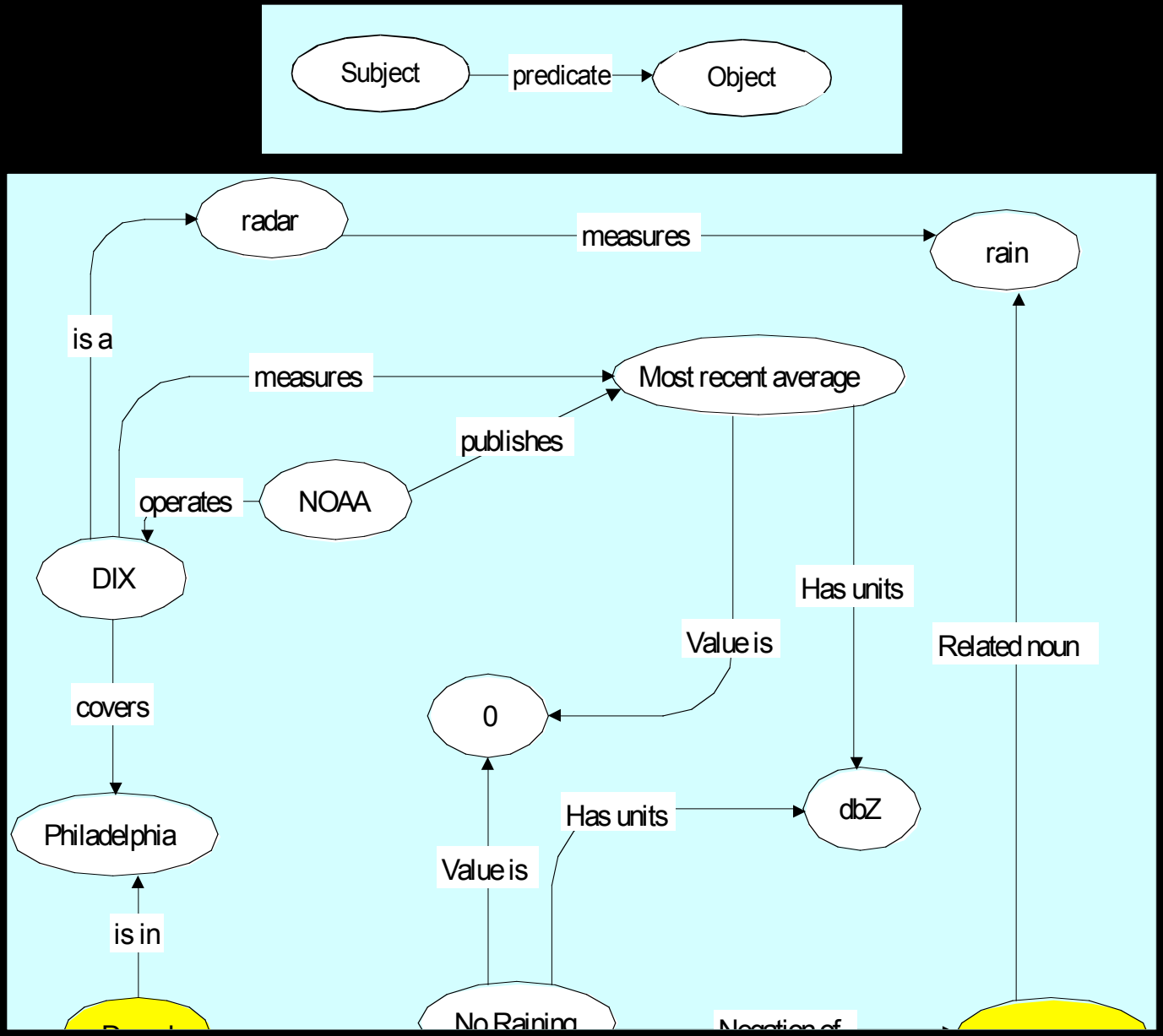
R Johnson (1998): Historical Research Online. A new ball game.

Data

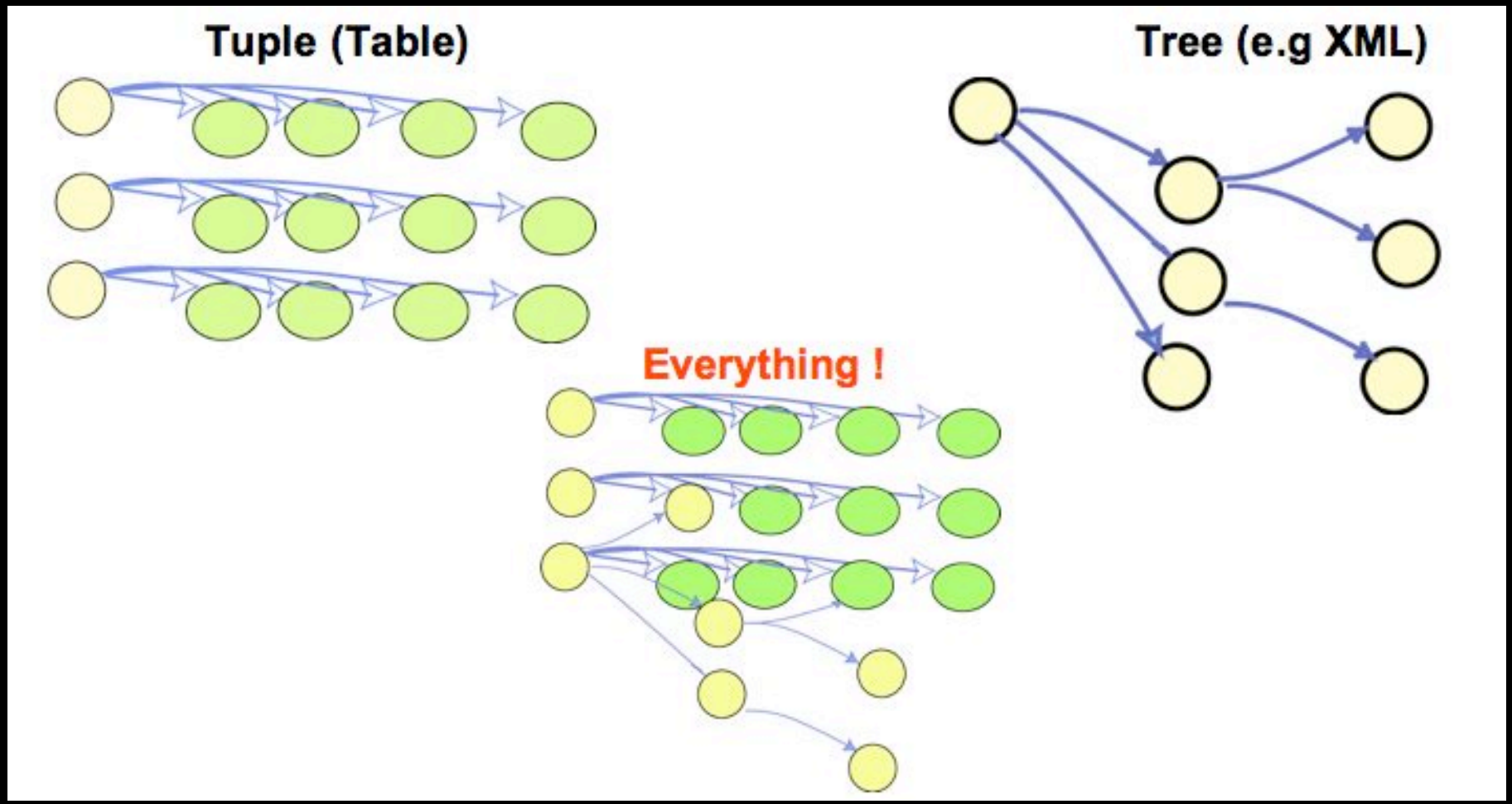
Information Resources (Metadata)

Knowledge Resources (RDF)

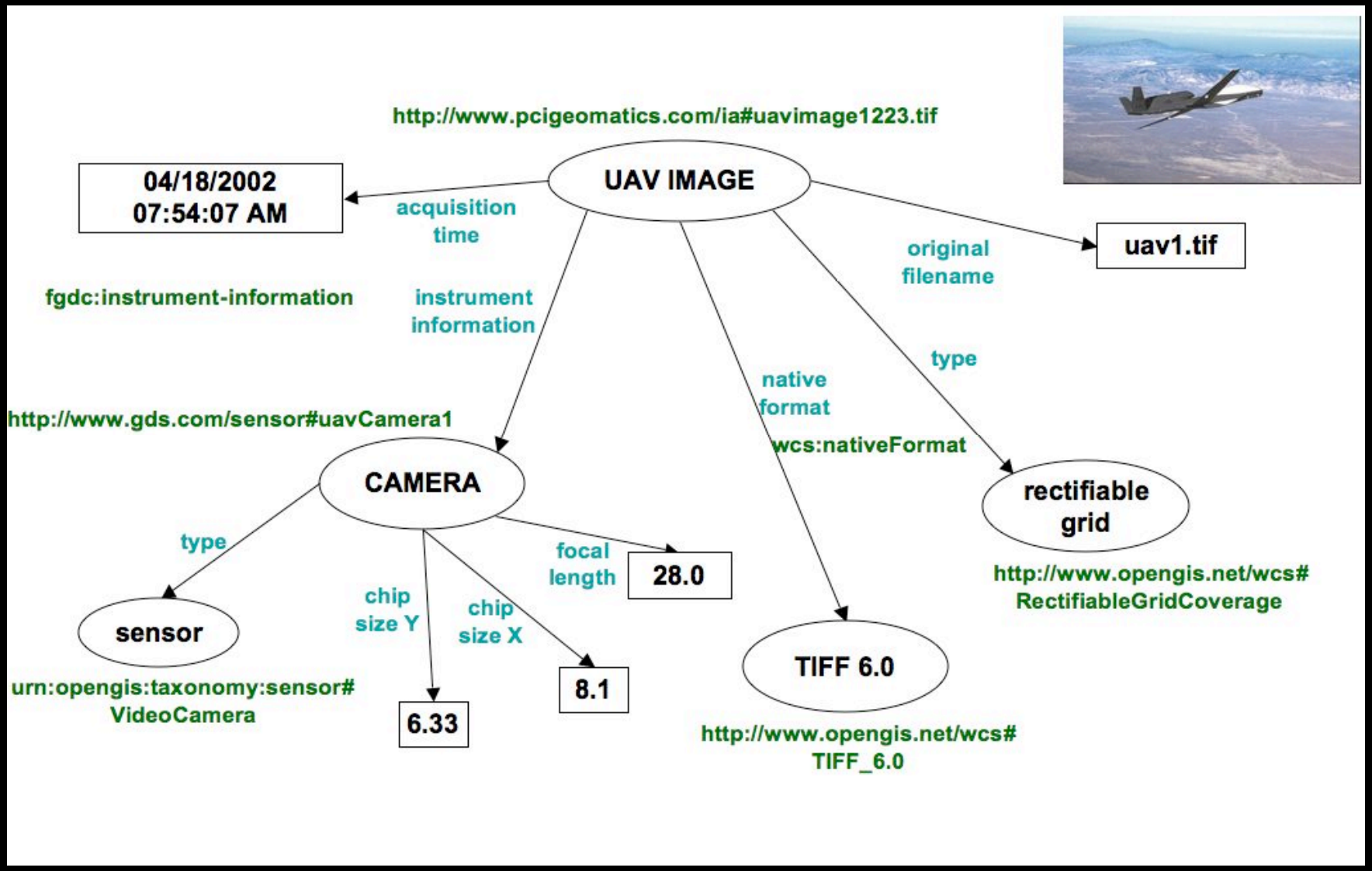
Resource Description Framework



RDF Simple Graph Model



Graphs = knowledge



URI

The most fundamental specification of Web architecture, while one of the simpler, is that of the URI. The principle that anything, absolutely anything, “on the web” should be identified distinctly by an otherwise opaque string of characters is core to the universality”. Tim Bernee Lee.

Proven paradigm: mail, IP, Web, SIN, Telephone number, P2P, banking transaction

Ontology definition

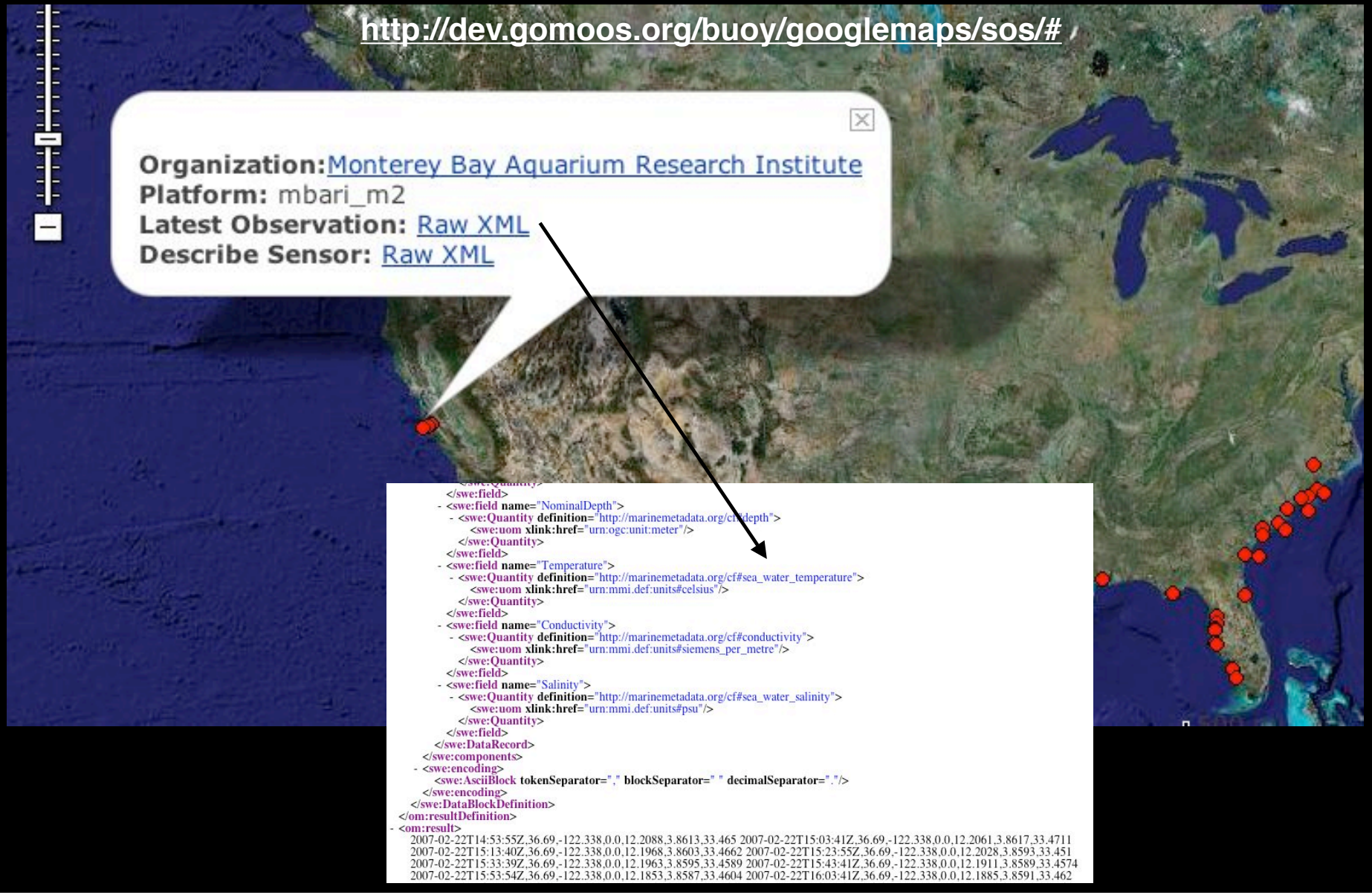
RDF Graphs which uses URIs as unique identifiers, useful to construct knowledge of domains, identifying, defining and relating control vocabularies.

Most Data and Metadata don't currently use URIs to identify semantics

Need to extract them. In other words, convert current controlled vocabularies to RDF, to enable reuse and graph construction in the Semantic World

MMI work:
CF, GCMD, UCUM, BODC, AGU, etc.
Some are in being created in realtime

Practical Application: OOSTETHYS



OOSTethys

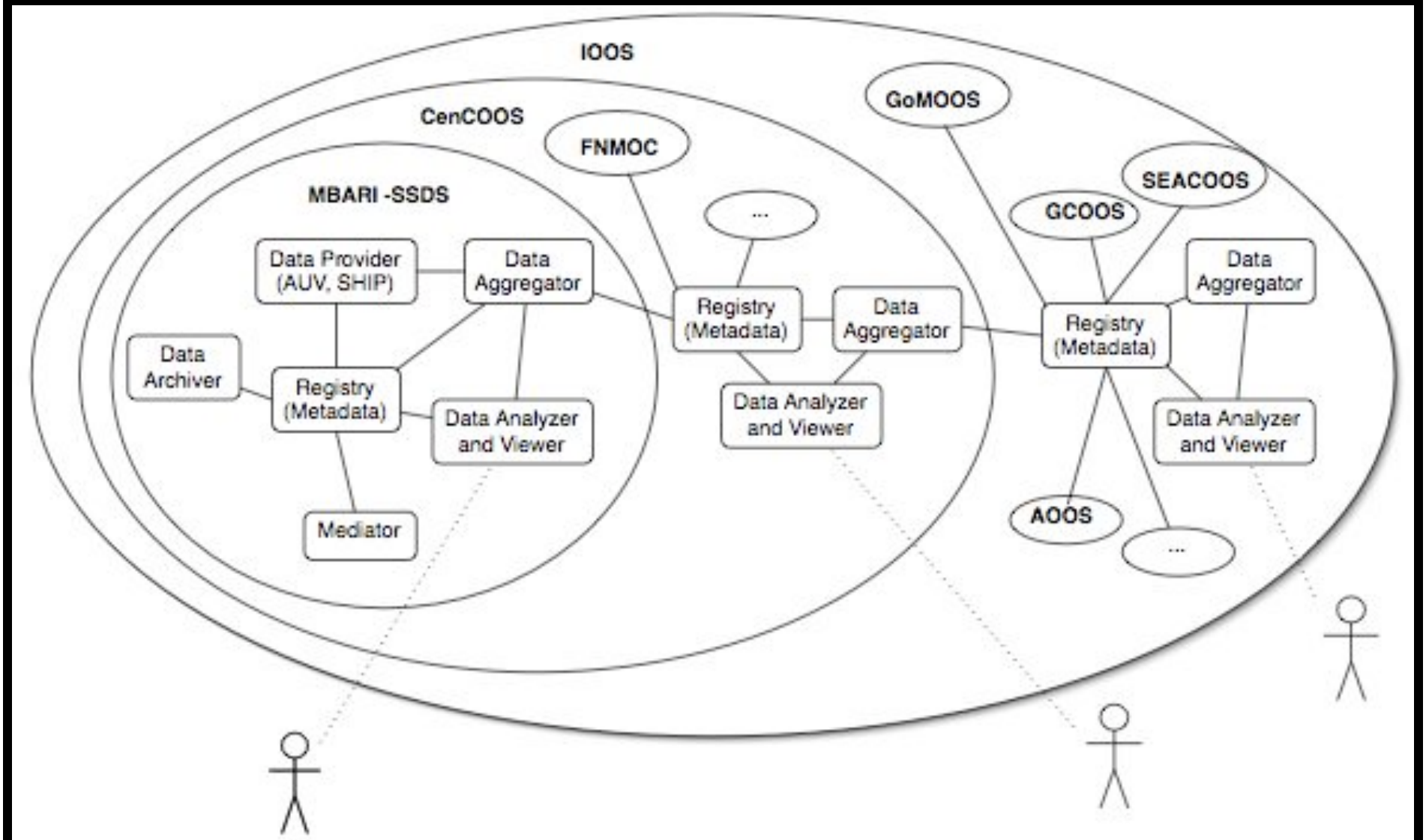
OOSTethys is an attempt to prototype a provider-to-user data systems framework, using interoperable standards, enabling discovery and use of data.

OOSTethys is currently focusing in enabling data providers with the software to implement an OGC-compliant Sensor Observation Service (SOS).

Participants

- Philip Bogden, GoMOOS & SURA
- Eric Bridger, GoMOOS
- David Forrest, VIMS
- Gerald Creager, Texas A&M
- John Graybeal , MBARI
- Jeremy Cothran, SEACOOS
- Helen Conover, UAH
- Bruce Beaumont, UAH
- Tony Cook, UAH
- Donna Cote, Texas A&M
- Luis Bermudez, MBARI
- Bill Howe, CMOP & NANOOS

Matryoshka


IOOS
CenCOOS
MBARI -SSDS
Data Provider (AUV, SHIP)
Data Aggregator
Data Archiver
Registry (Metadata)
Data Analyzer and Viewer
Mediator
FNMOC
...
Registry (Metadata)
Data Aggregator
Data Analyzer and Viewer
GoMOOS
SEACOOS
GCOOS
Data Aggregator
Registry (Metadata)
Data Analyzer and Viewer
AOOS
...

Sensor Observation Service (SOS) Main Interfaces

getCapabilities → list of Sensors (Sensor ID)

list of Observation Offerings (OOId)

- PhenomenaIds
- ProcessesIds
- FutureOfInterestIds

describeSensor (Sensor ID) → Description, physical characteristics, interfaces, inputs, outputs, lineage, constraints, calibration and accuracy.

getObservation (OOId) → Data

A lot of Ids ! which need URIs

Vocabularies Agreement

Namespaces:

- cf = http://marinemetadata.org/cf#
- gcmd = http://marinemetadata.org/gcmd#

Observed Properties

Component definitions and units of measurement

Common terms	Data producers terms (components definition)	Units of measure	Portal terms (discovery – qualifiers)
time	urn:ogc:phenomenon:time:iso8601	not necessary	urn:ogc:phenomenon:time:iso8601
latitude	urn:ogc:phenomenon:latitude:wgs84	urn:ogc:unit:degree	urn:ogc:phenomenon:latitude:wgs84
longitude	urn:ogc:phenomenon:longitude:wgs84	urn:ogc:unit:degree	urn:ogc:phenomenon:longitude:wgs84
salinity, SSS, sea water salinity, sea surface salinity	cf:sea_water_salinity, cf:sea_surface_salinity	urn:mmi.def:units#psu	gcmd:Oceans--Salinity_Density--Salinity
depth	cf:depth	urn:ogc:unit:meter	gcmd:Oceans--Surface_Water--Water_Depth

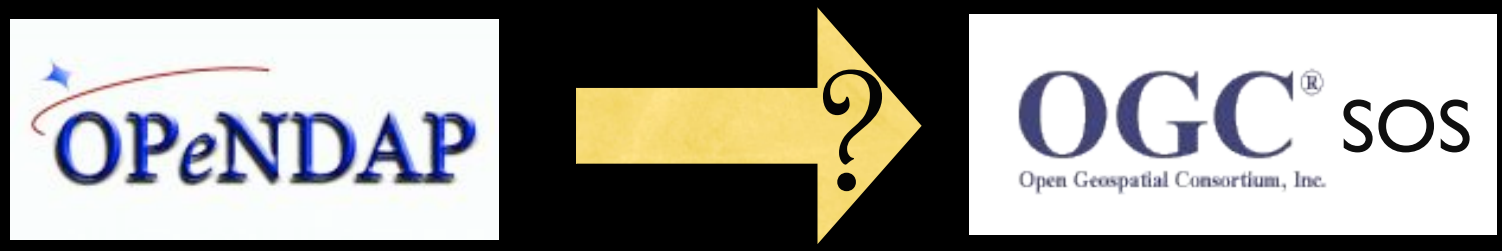
This can store in an ontology !

OOSTethys Cookbooks

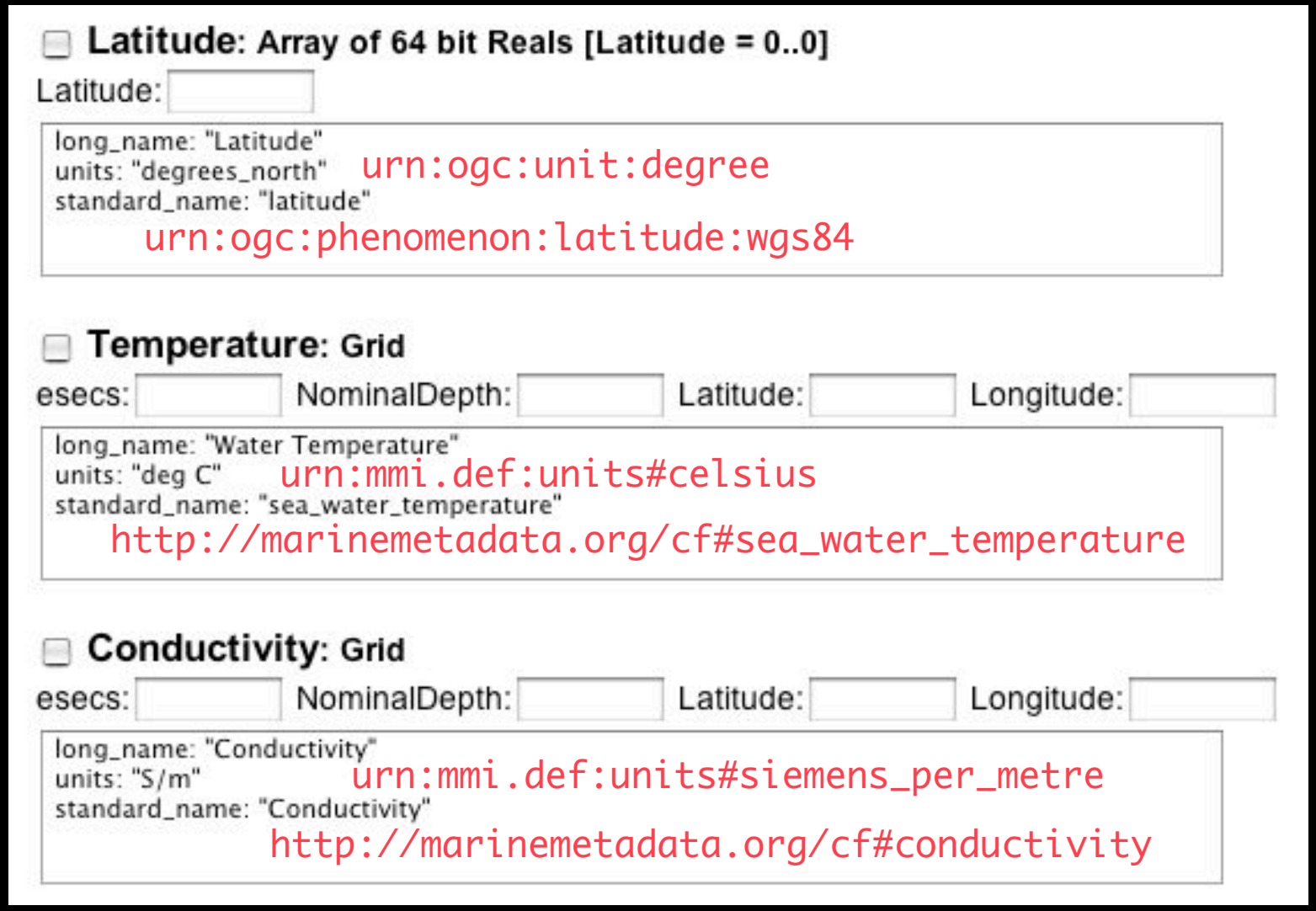
OOSTethys allows a data provider to setup an OGC Sensor Observation Service (SOS) with minimal effort, by publishing cookbooks, templates and making available source code. Implementation examples:

- PERL from ASCII files
- Python from RDBS
- JAVA from OpenDAP

OPeNDAP to SOS



Need to create semantics (URIs) from OPenDAP metadata



Need to create semantics (URIs) from OPenDAP metadata

Need more than just variables. For example:

Process Type (e.g. Sensor or Platform Id)

Feature of Interest (e.g. water body name)

So...

From this OPeNDAP uris

- http://dods.mbari.org/cgi-bin/nph-nc/data/ssdsdata/deployments/m0/current_netCDFs/ctd0000.nc
- http://dods.mbari.org/cgi-bin/nph-nc/data/ssdsdata/deployments/m1/current_netCDFs/ctd0000.nc
- http://dods.mbari.org/cgi-bin/nph-nc/data/ssdsdata/deployments/m2/current_netCDFs/ctd0000.nc

An SOS is created. The capabilities document can be found here:

http://marinemetadata.org/mmiws/oostethys/sos?VERSION=0.0.31&SERVICE=SOS&REQUEST=GetCapabilities

[for copy paste pleaseee]

Example Observation Offering

```
- <sos:ObservationOffering gml:id="mbari_m2">
    <gml:description/>
    <gml:name/>
  - <gml:boundedBy>
    - <gml:Envelope>
        <gml:lowerCorner srsName="urn:ogc:def:crs:EPSG:6.5:4329">36.69 -122.338 0</gml:lowerCorner>
        <gml:upperCorner srsName="urn:ogc:def:crs:EPSG:6.5:4329">36.69 -122.338 0</gml:upperCorner>
      </gml:Envelope>
    </gml:boundedBy>
  - <sos:time>
    - <gml:TimePeriod gml:id="mbari_m2_offeringTime">
        <gml:beginPosition>2006-03-30T20:30:13Z</gml:beginPosition>
        <gml:endPosition>2007-02-22T16:03:41Z</gml:endPosition>
      </gml:TimePeriod>
    </sos:time>
    <sos:procedure xlink:href="urn:org:mbari:mooring#M2"/>
    <sos:observedProperty xlink:href="http://marinemetadata.org/cf#sea_water_temperature"/>
    <sos:observedProperty xlink:href="http://marinemetadata.org/cf#conductivity"/>
    <sos:observedProperty xlink:href="http://marinemetadata.org/cf#sea_water_salinity"/>
    <sos:featureOfInterest xlink:href="urn:mmi.feature#bodyOfWater"/>
    <sos:responseFormat>application/com-xml</sos:responseFormat>
```

Code available

http://www.oostethys.org

Wrap up slide

- Enabling semantics is adding knowledge
- Currently Semantic Web tools rely on RDF (Simple graph)
- URIs increasingly popular (this includes OGC folks)
- Need more URIs, but managing ontologies not easy (~ ask in the CF list how hard is to make agreements)
- Ontologies are and "should always be" a community agreement
- MMI is here to help - one of the main goals is to build community and participate in interoperability demonstrations - e.g OOSTETHYS

more...

- OGC Ocean Science Interoperability Experiment has been released:

  http://www.opengeospatial.org/projects/initiatives/oceansie

  Thank you...

  Luis Bermudez (bermudez@mbari.org)