OCM Ontology and Ontology Services

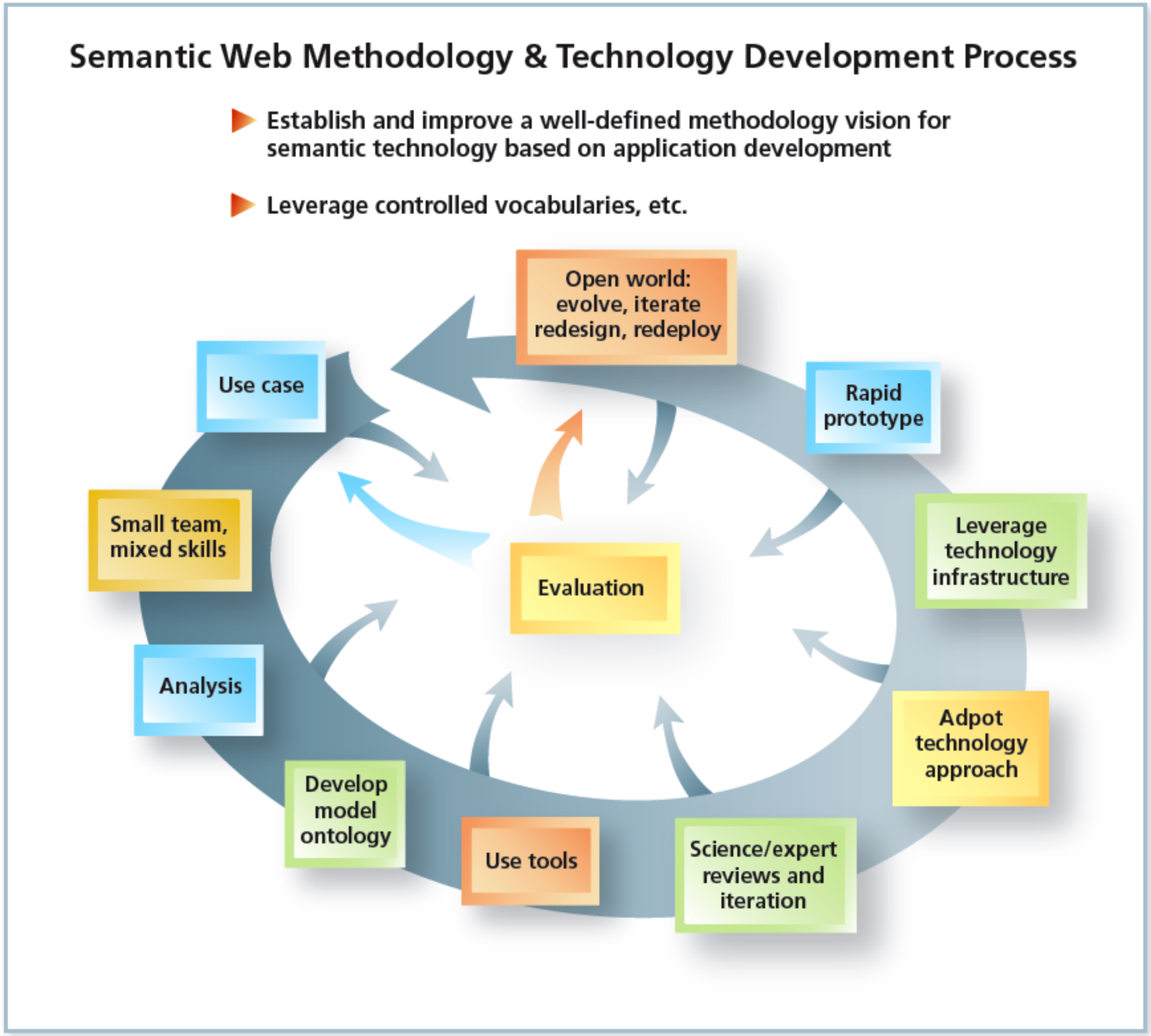
August 14, 2012

NOAA, Boulder CO

Peter Fox (RPI* and WHOI**) pfox@cs.rpi.edu and OTHERS!!
*Tetherless World Constellation, ** AOP&E

Modern informatics

- Use cases
- Stakeholders
- Distributed authority
- Access control
- Ontologies
- Maintaining Identity



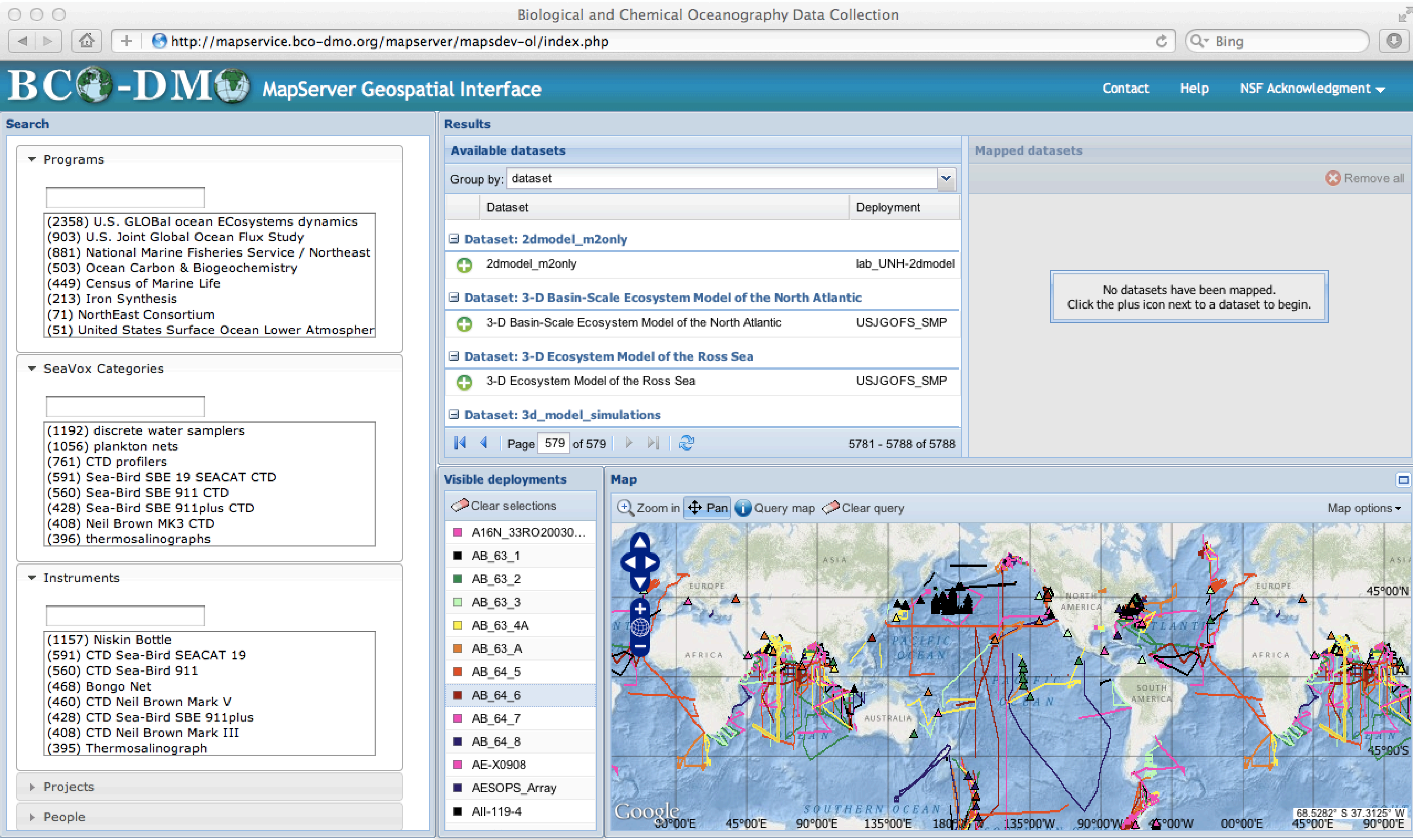
Use Case

- … is a collection of possible sequences of interactions between the system under discussion and its actors, relating to a particular goal.
- The collection of Use Cases should define all system behavior relevant to the actors to assure them that their goals will be carried out properly.
  - is a prose description of a system's behavior when interacting with the outside world.
  - is a technique for capturing functional requirements of business systems and, potentially, of an ICT system to support the *business* system.
  - can also capture non-functional requirements

IOCM Use Case

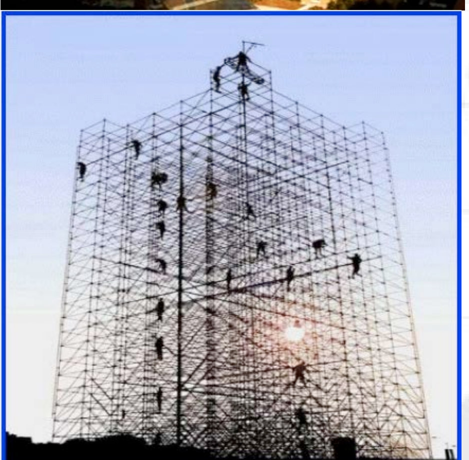
- “Provide a list of NOAA datasets in a particular geographic area including quality, extent, responsible agency and datatype with the intent of justifying NOAA contributions to the IOCM.”

What might an implementation look like?



Systems v. Frameworks

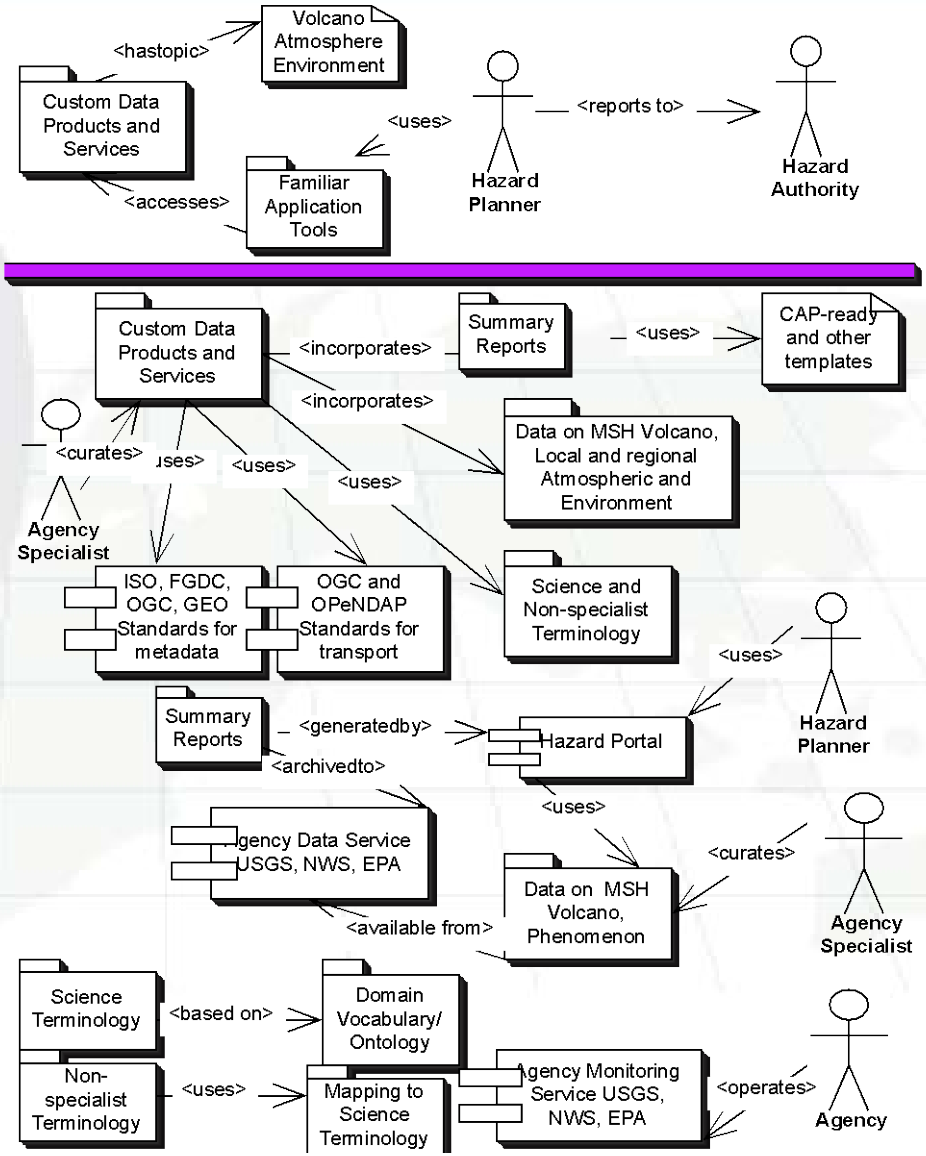
- Rough definitions



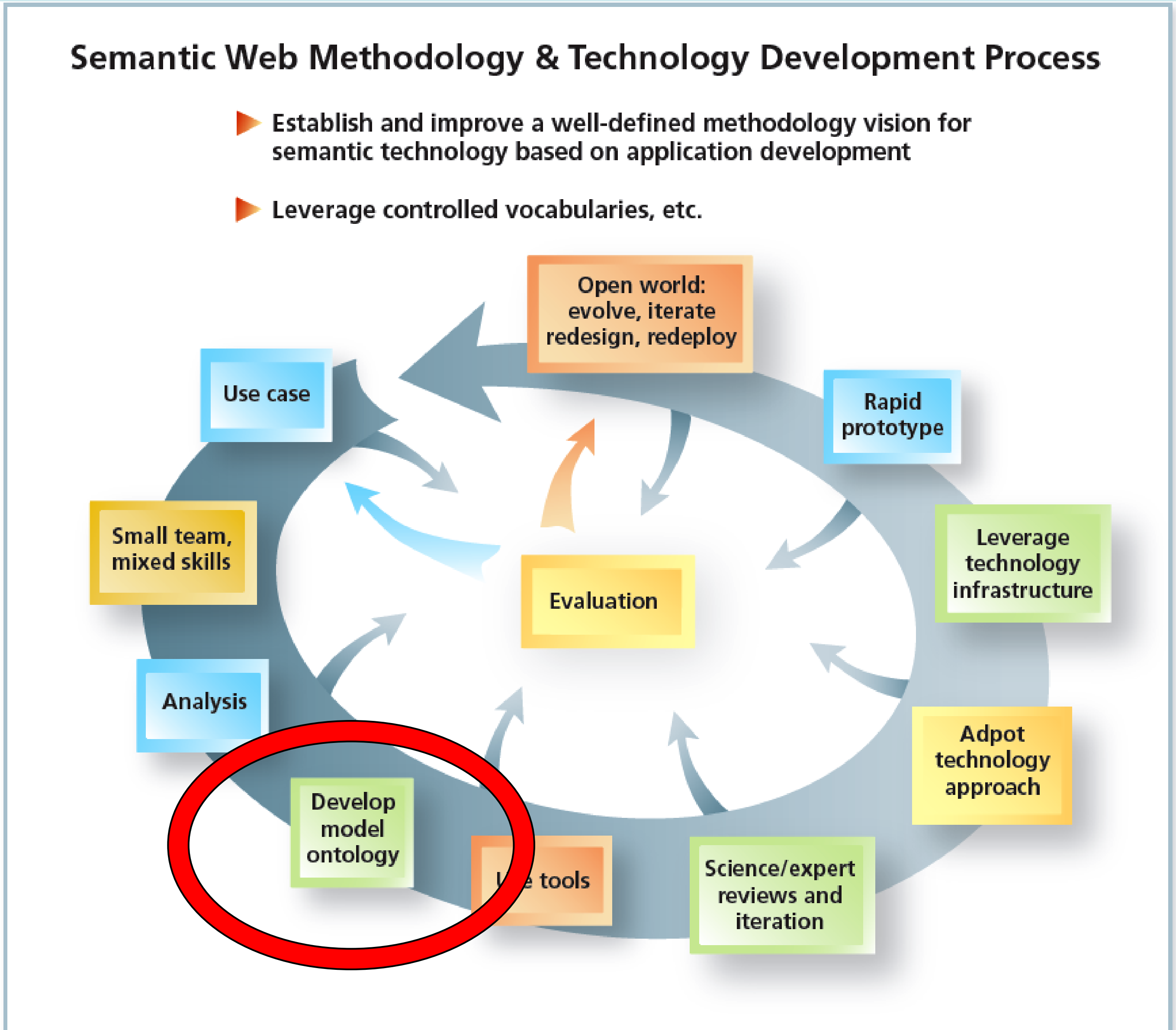
  - Systems have very well-define entry and exit points. A user tends to know when they are using one. Options for extensions are limited and usually require engineering
  - Frameworks have many entry and use points. A user often does not know when they are using one. Extension points are part of the design
  - Platforms are built on frameworks

Use Cases Expose System Requirements

- Exposes goals, outcomes, actors/ roles, resources, preconditions, process flow, artifacts
- And … semantics, terms, concepts and their relations



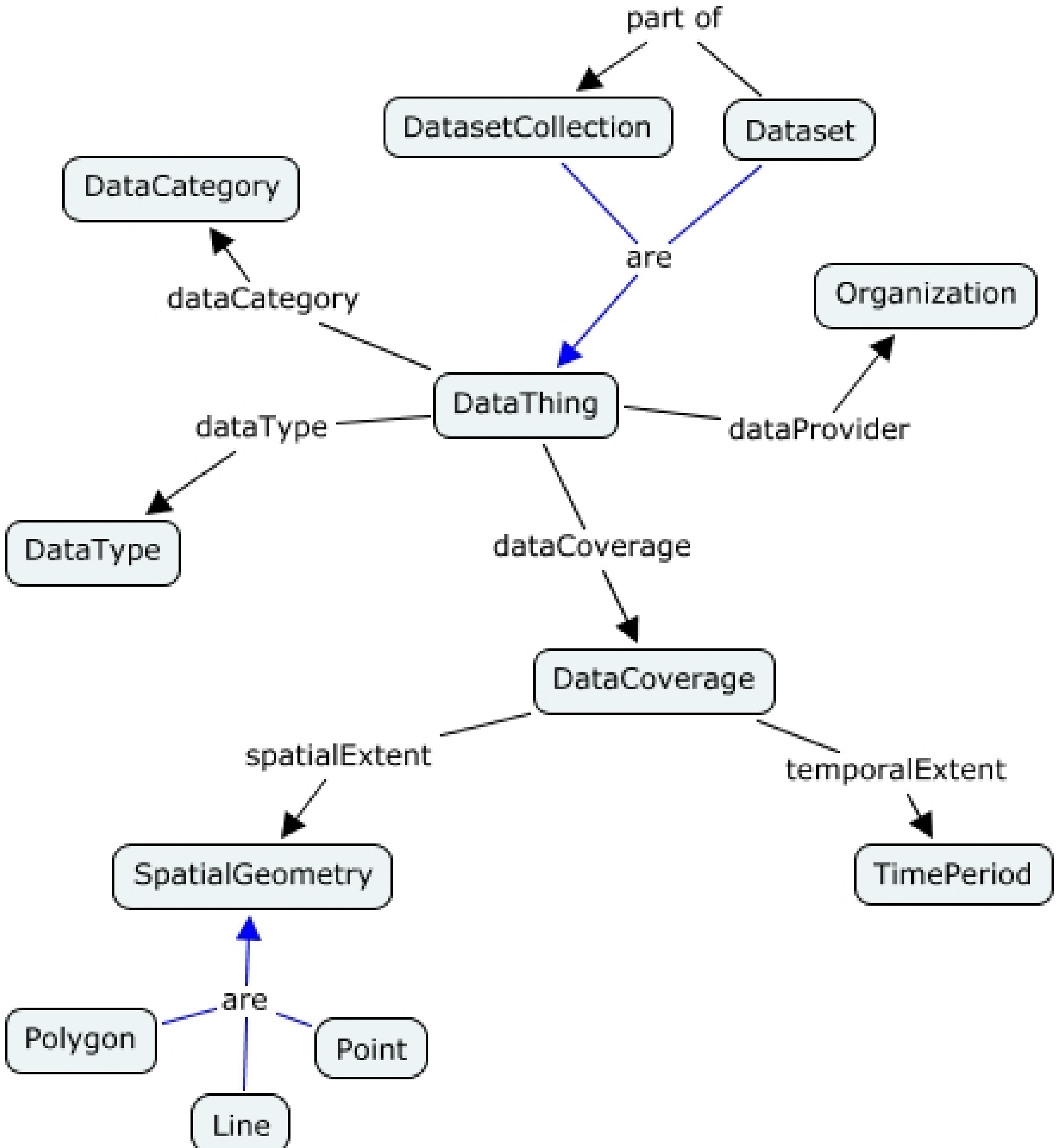
Model


Semantic Web Methodology & Technology Development Process
Establish and improve a well-defined methodology vision for semantic technology based on application development
Leverage controlled vocabularies, etc.
Open world: evolve, iterate redesign, redeploy
Use case
Rapid prototype
Small team, mixed skills
Leverage technology infrastructure
Evaluation
Analysis
Adpot technology approach
Develop model ontology
tools
Science/expert reviews and iteration

Information modeling

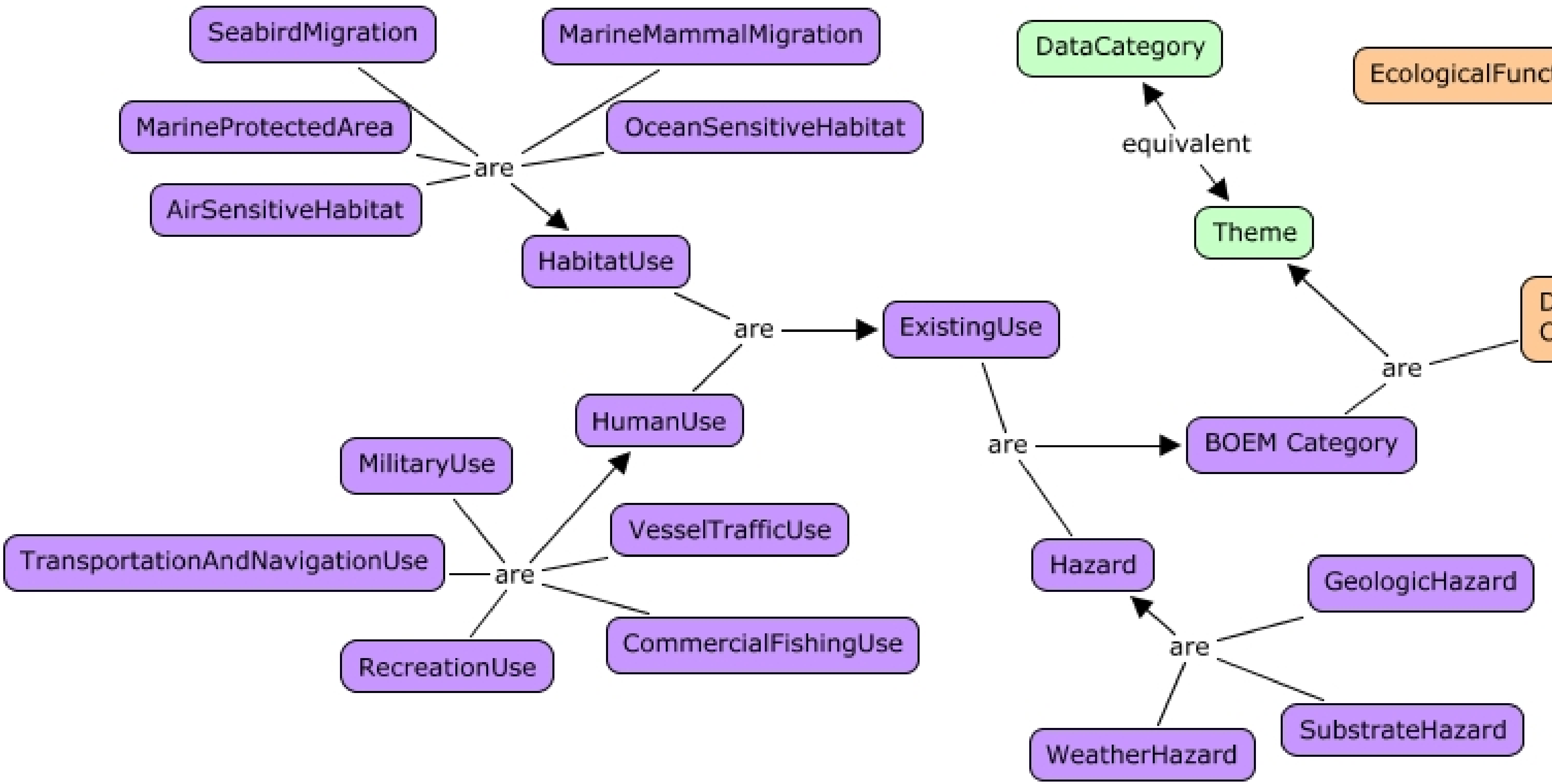
- Conceptual models, sometimes called domain models, are typically used to explore domain concepts and often created
  - as part of initial requirements envisioning efforts as they are used to explore the high-level static *business or science or medicine* structures and concepts
  - as the precursor to logical models or as alternatives to them
- Followed by logical and physical models
- Introduced in ANSI processes in 1978!

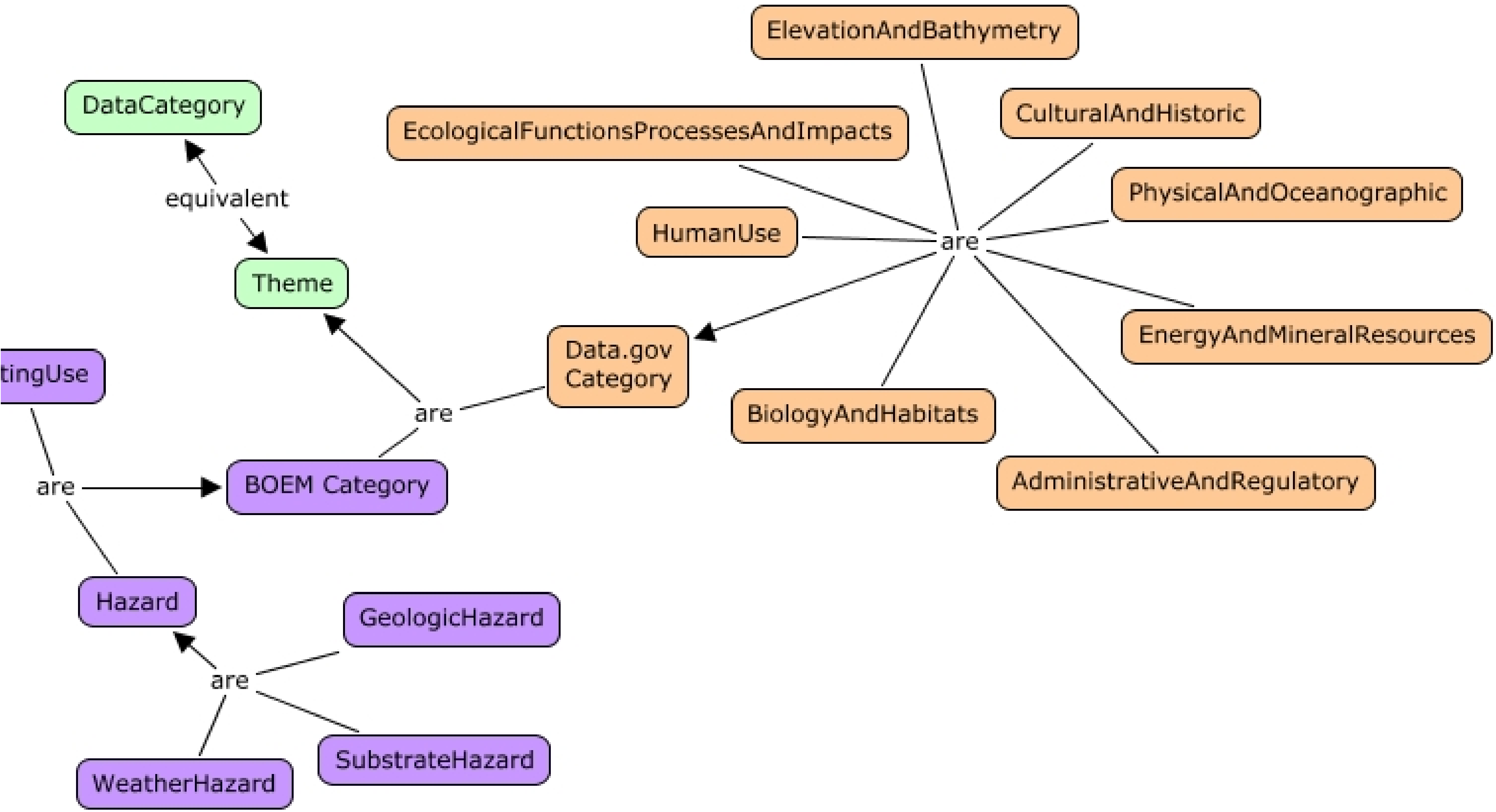
DataType and others…



Another use case

- To identify priority wind energy areas for potential development and accelerate the leasing process by evaluating existing ocean uses by humans and natural resources.
- Coastal and Marine Spatial Planning


SeabirdMigration
MarineMammalMigration
MarineProtectedArea
OceanSensitiveHabitat
AirSensitiveHabitat
are
HabitatUse
are
ExistingUse
HumanUse
MilitaryUse
VesselTrafficUse
TransportationAndNavigationUse
are
CommercialFishingUse
RecreationUse
DataCategory
equivalent
Theme
EcologicalFunc
are
are
BOEM Category
Hazard
GeologicHazard
are
SubstrateHazard
WeatherHazard



ElevationAndBathymetry
DataCategory
EcologicalFunctionsProcessesAndImpacts
CulturalAndHistoric
PhysicalAndOceanographic
equivalent
HumanUse
are
Theme
Data.gov Category
EnergyAndMineralResources
tingUse
are
BiologyAndHabitats
AdministrativeAndRegulatory
are
BOEM Category
Hazard
GeologicHazard
are
SubstrateHazard
WeatherHazard

Logical models

- For a logical data model to be normalized, it must include the **full population of attributes** to be implemented and those attributes must be defined in terms of their domains or **logical data types** (e.g., character, number, date, picture, etc.).
- A logical data model requires a complete scheme of **identifiers** or candidate keys for unique identification of each occurrence in every entity

- A *logical* model is provable in the mathematics of data science. E.g. for relational databases, logical models generally conform to relational theory.
- Thus a logical model contains only **fully normalized entities**. Some of these may represent logical domains rather than, for e.g. potential physical tables.

Object oriented design

- Object-oriented modeling is a formal way of representing something in the real world (draws from traditional set theory and classification theory). Some basics to keep in mind in object-oriented modeling are that:
  - **Instances are *things.***
  - ***Properties are attributes.***
  - ***Relationships are pairs of attributes.***
  - ***Classes are types of things.***
  - ***Subclasses are subtypes of things.***

Physical models

- A *physical* model is a single logical model instantiated in a specific information system (e.g., relational database, RDF/XML document, etc.) in a specific installation.
- The physical model specifies implementation details which may be features of a particular product or version, as well as configuration choices for that instance.

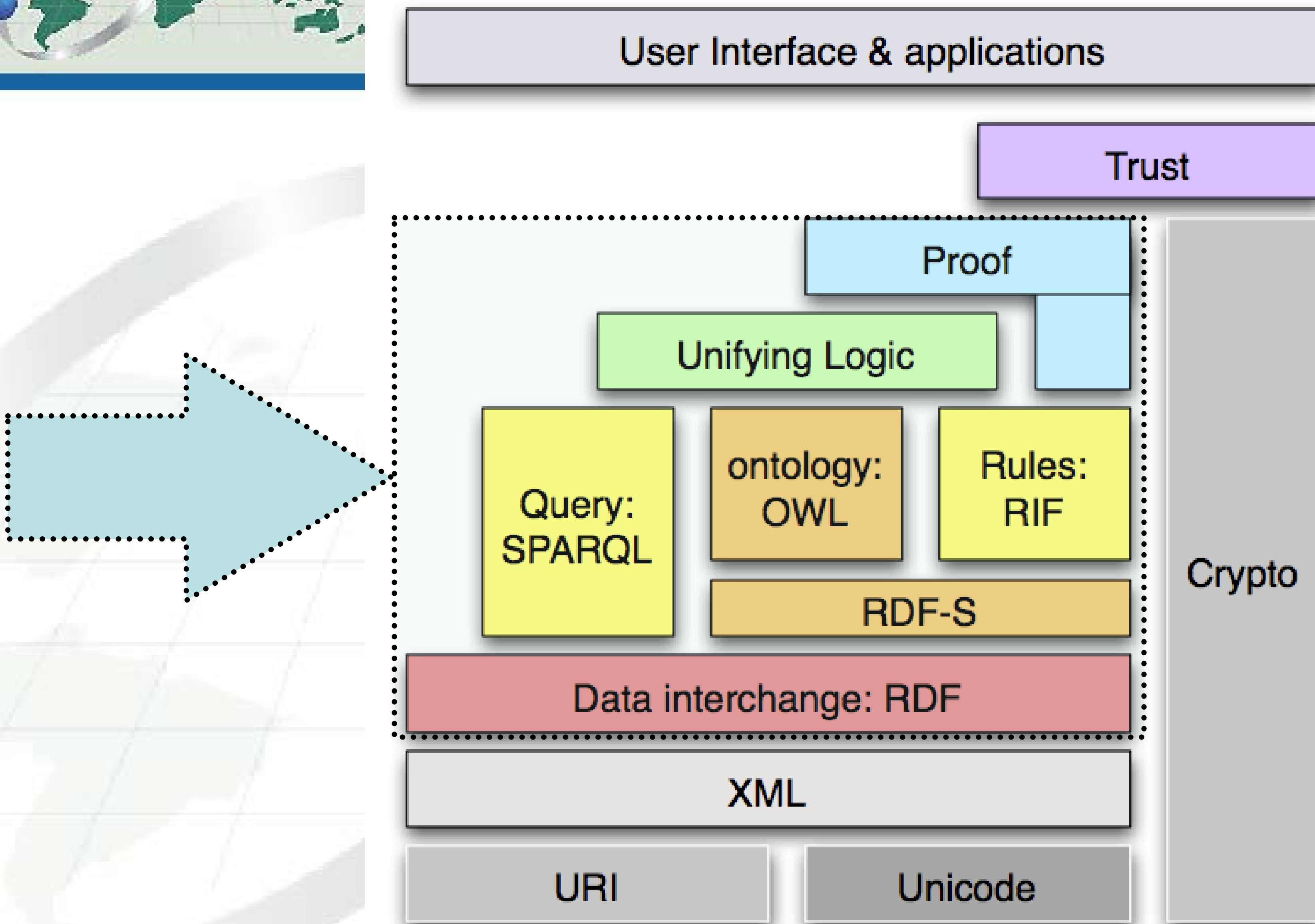
For example for relational DBs

Feature	Conceptual	Logical	Physical
Entity Names	✓	✓	
Entity Relationships	✓	✓	
Attributes		✓	
Primary Keys		✓	✓
Foreign Keys		✓	✓
Table Names			✓
Column Names			✓
Column Data Types			✓

Not an isolated set of models

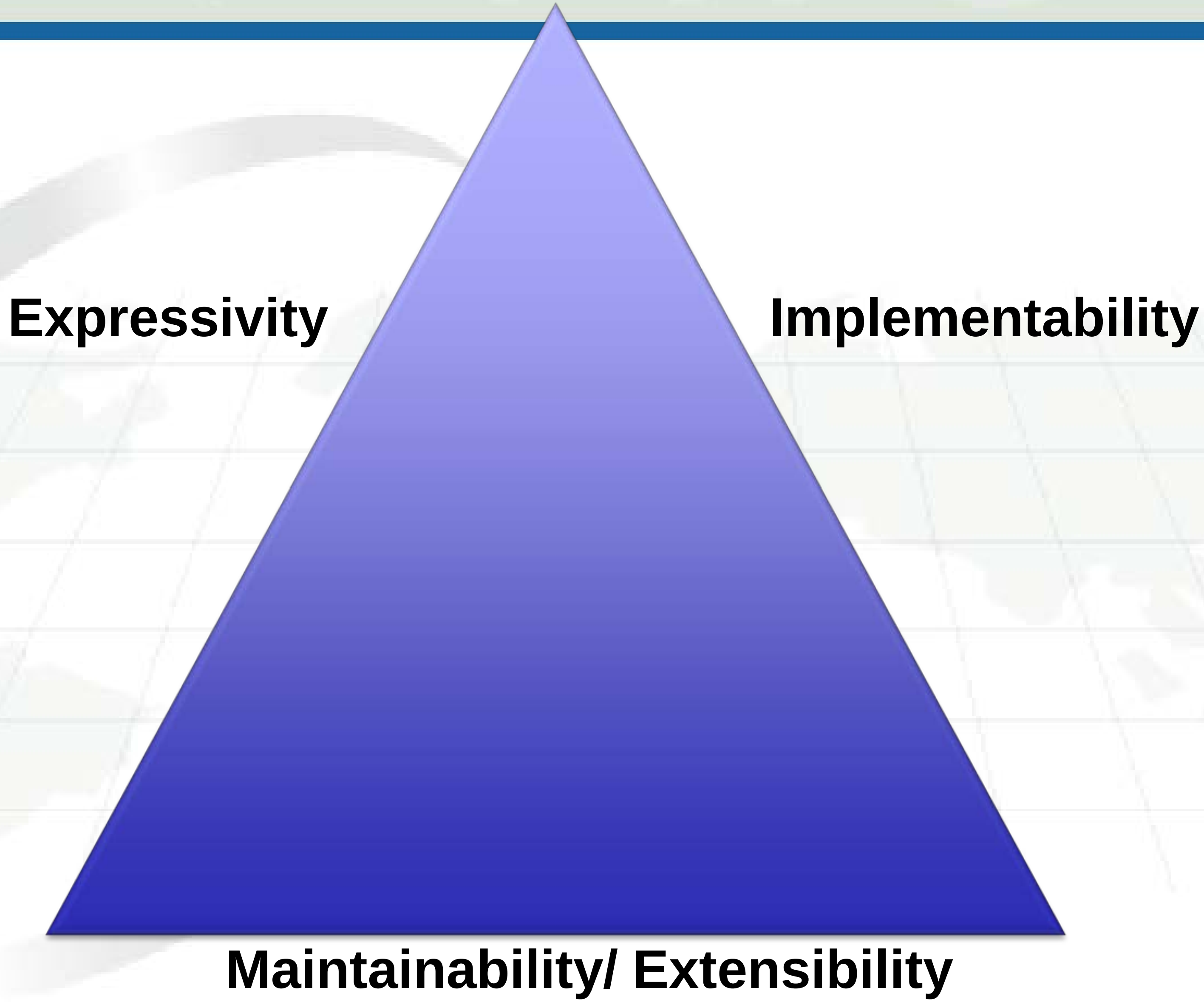
- Handle errors, iteration, evolution, …
  - To the logical model?
  - To the conceptual model?
- Relating to and/ or integrating with other information models?
  - General rule – integrate at the highest level you can (i.e. more abstract)

Semantic Web Layers

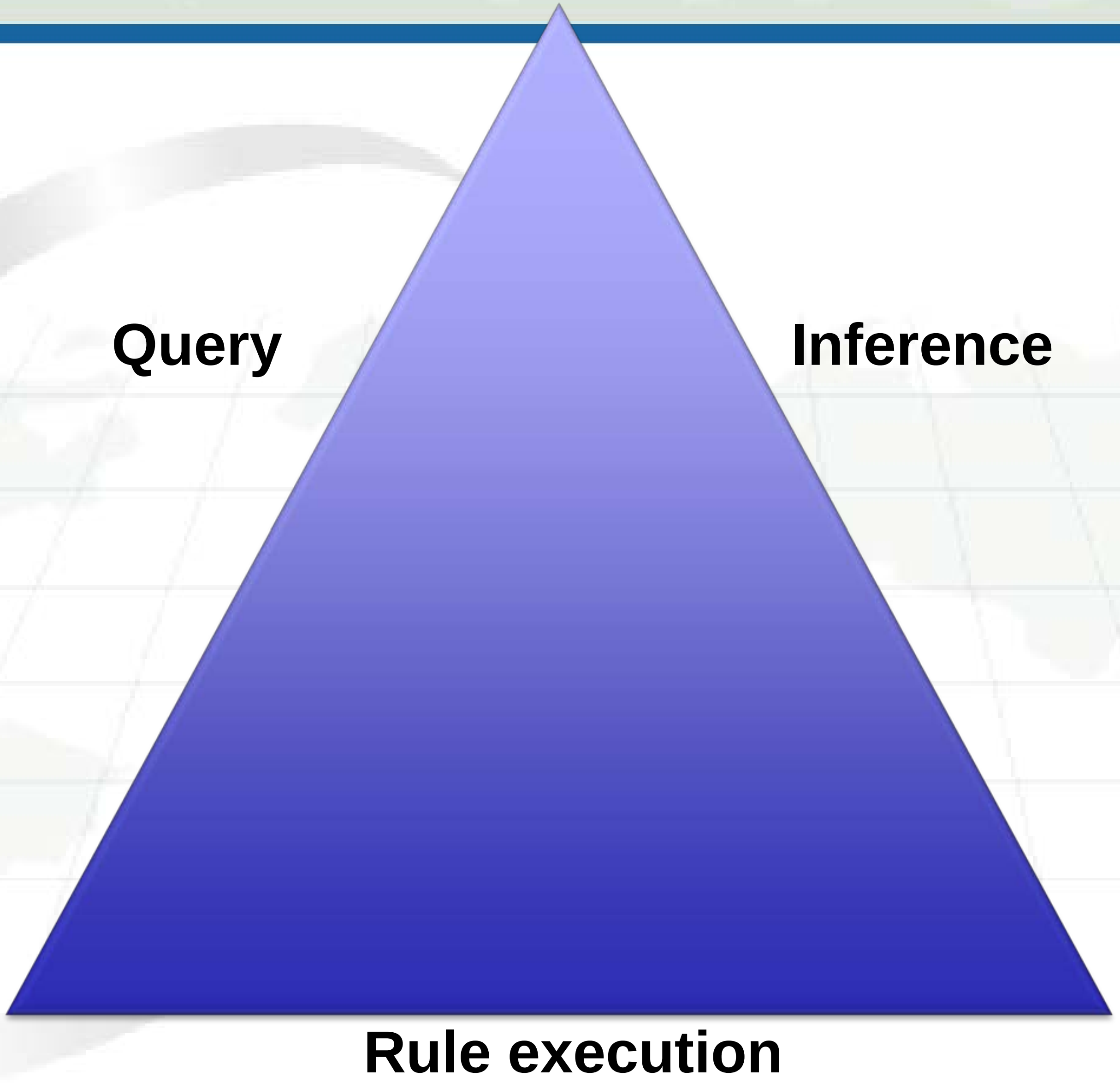


http://www.w3.org/2003/Talks/1023-iswc-tbl/slide26-0.html, http://flickr.com/photos/pshab/291147522/

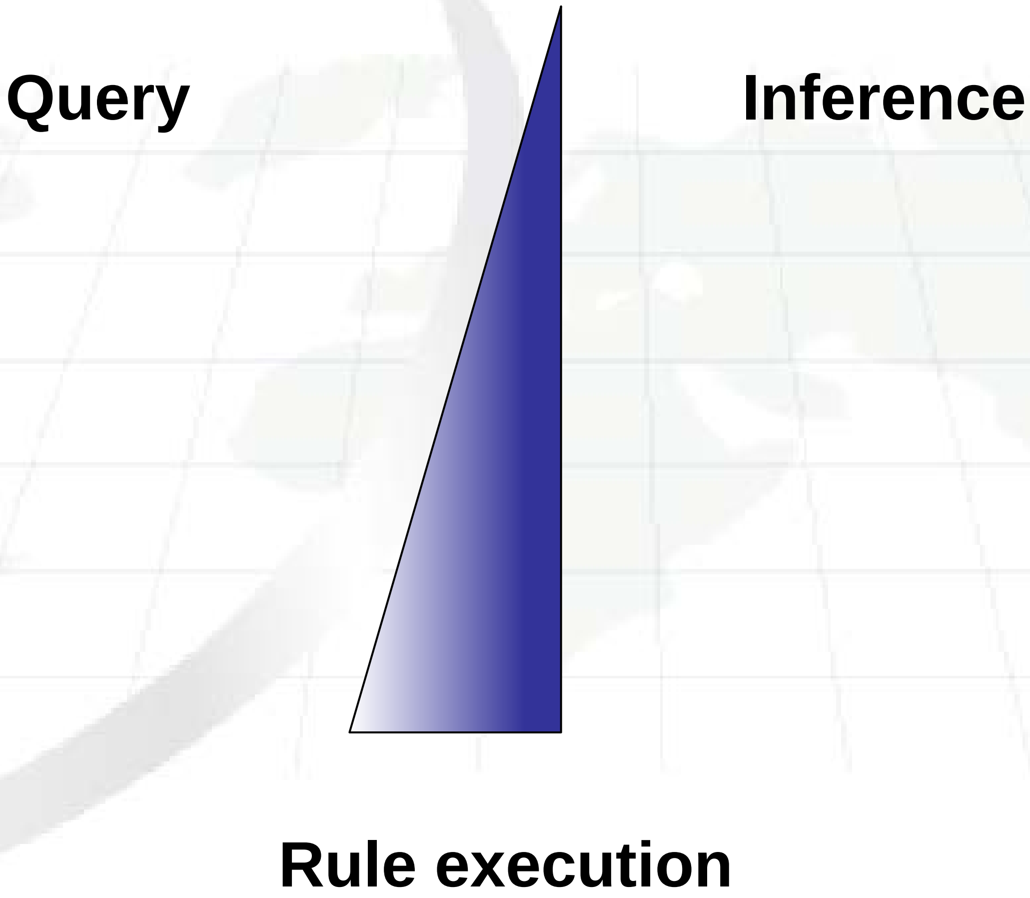
Working with knowledge



Working with knowledge



Or it may be this …



Expressivity/ Implementation

Declarative

Procedural

Linked open data
URI/http/RDF *

Ontology encoded

Semantic Web Standards*

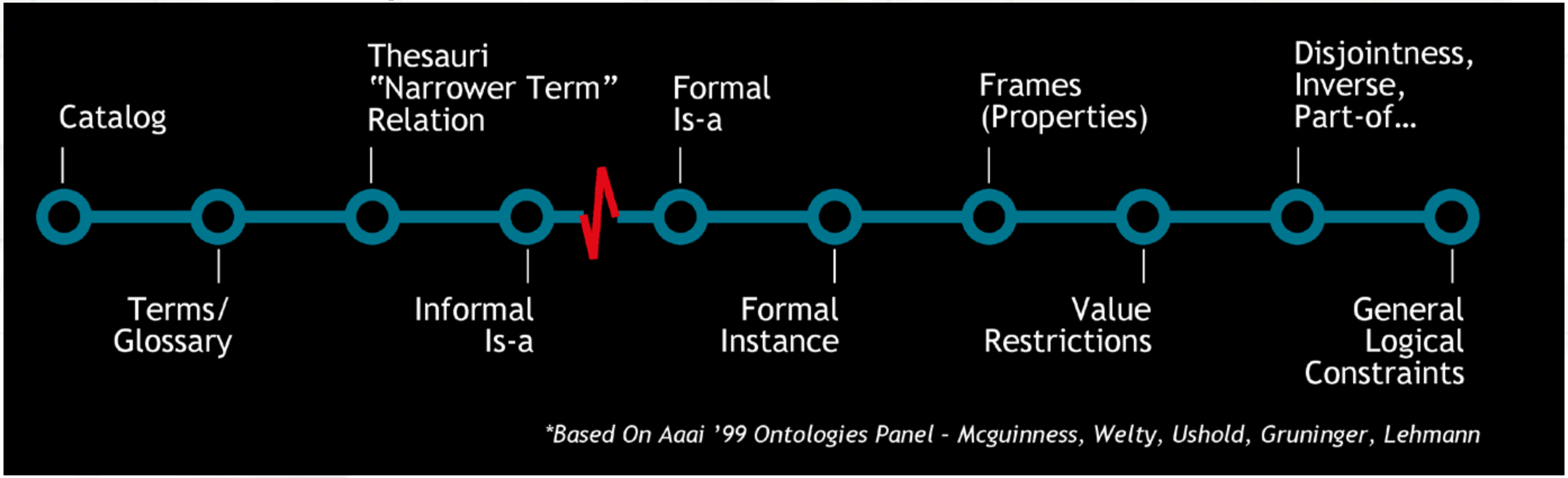
- Schema - RDFS (Resource Description Framework Schema, 2004)
- Ontology - OWL 1.0 (Web Ontology Language, 2004)
- Query - SPARQL 1.0 (**S**PARQL **P**rotocol **a**nd **R**DF **Q**uery **L**anguage, 2008)
- OWL 2.0 (2009)
- Taxonomy - SKOS (Simple Knowledge Organization System, 2009)
- Rules - RIF (Rule Interchange Framework, 2010)
- SPARQL 1.1 (in review)
- NB. No service standards!

Ontology Spectrum

An ontology specifies a rich description of the

- Terminology, concepts, nomenclature
- Properties explicitly defining concepts
- Relations among concepts (hierarchical and lattice)
- Rules distinguishing concepts, refining definitions and relations (constraints, restrictions, regular expressions)

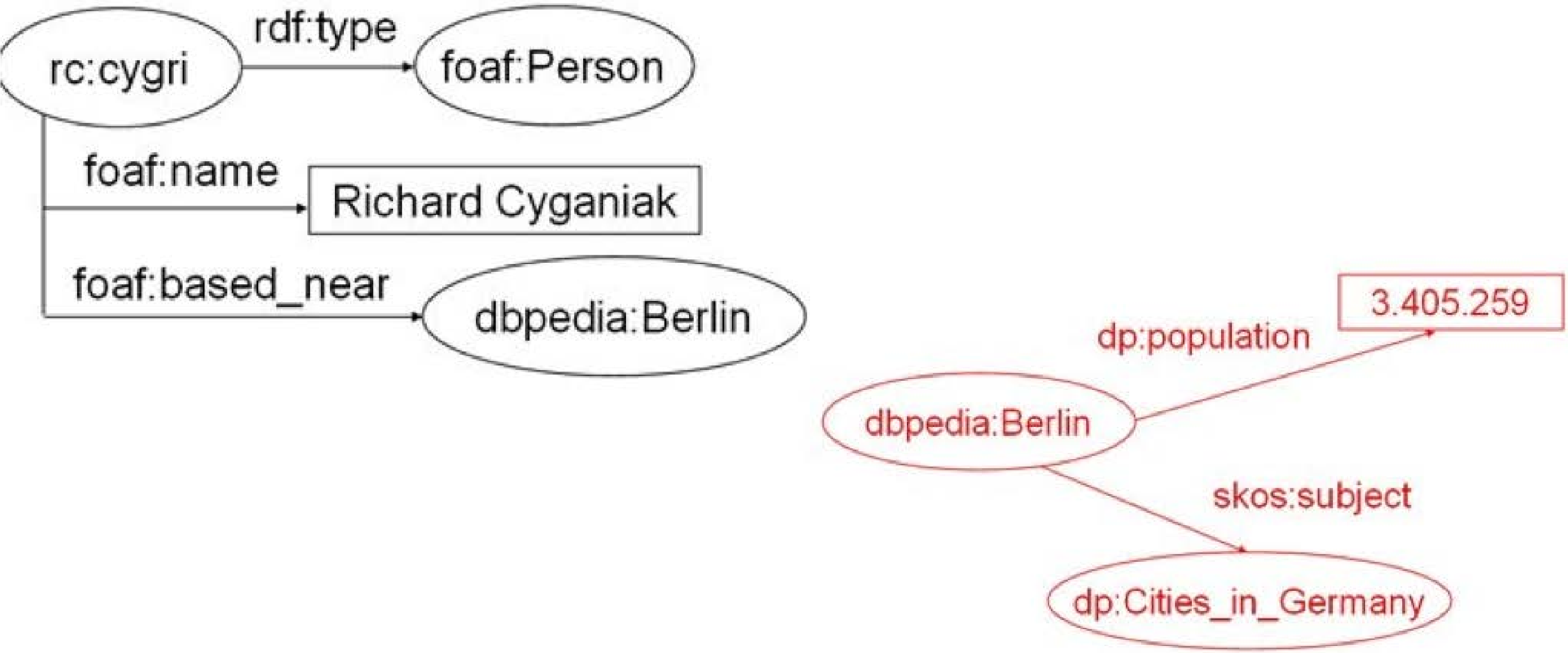
relevant to a particular domain or area of interest.



www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-abstract.html

slide from Kendall/McGuinness SemTech Tutorial

Unit of exchange – the *triple* - example (linked data)



Heath (2009)

RDFS

- Note: XMLS *not* an ontology language
  - Changes format of DTDs (document schemas) to be XML
  - Adds an extensible type hierarchy
    - Integers, Strings, etc.
    - Can define sub-types, e.g., positive integers
- RDFS *is* recognisable as an ontology language
  - Classes and properties
  - Sub/super-classes (and properties)
  - Range and domain (of properties)

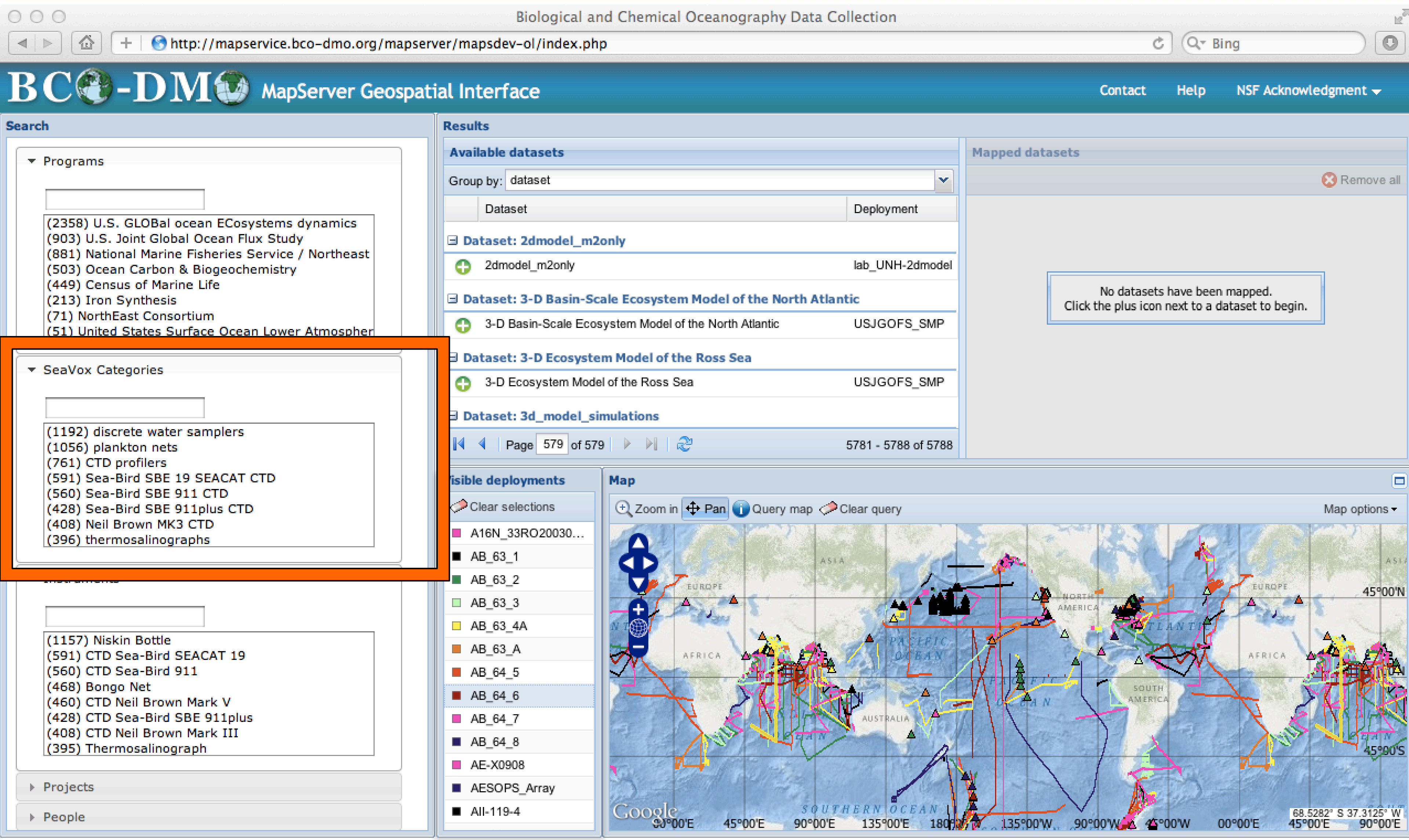
However

- RDFS too weak to describe resources in sufficient detail
  - No localized range and domain constraints
    - Can't say that the range of hasChild is person when applied to persons and elephant when applied to elephants
  - No existence/cardinality constraints
    - Can't say that all *instances* of person have a mother that is also a person, or that persons have exactly 2 parents
  - No transitive, inverse or symmetrical properties
    - Can't say that isPartOf is a transitive property, that hasPart is the inverse of isPartOf or that touches is symmetrica
- Difficult to provide reasoning support
  - No "native" reasoners for non-standard semantics
  - May be possible to reason via First Order axiomatisation

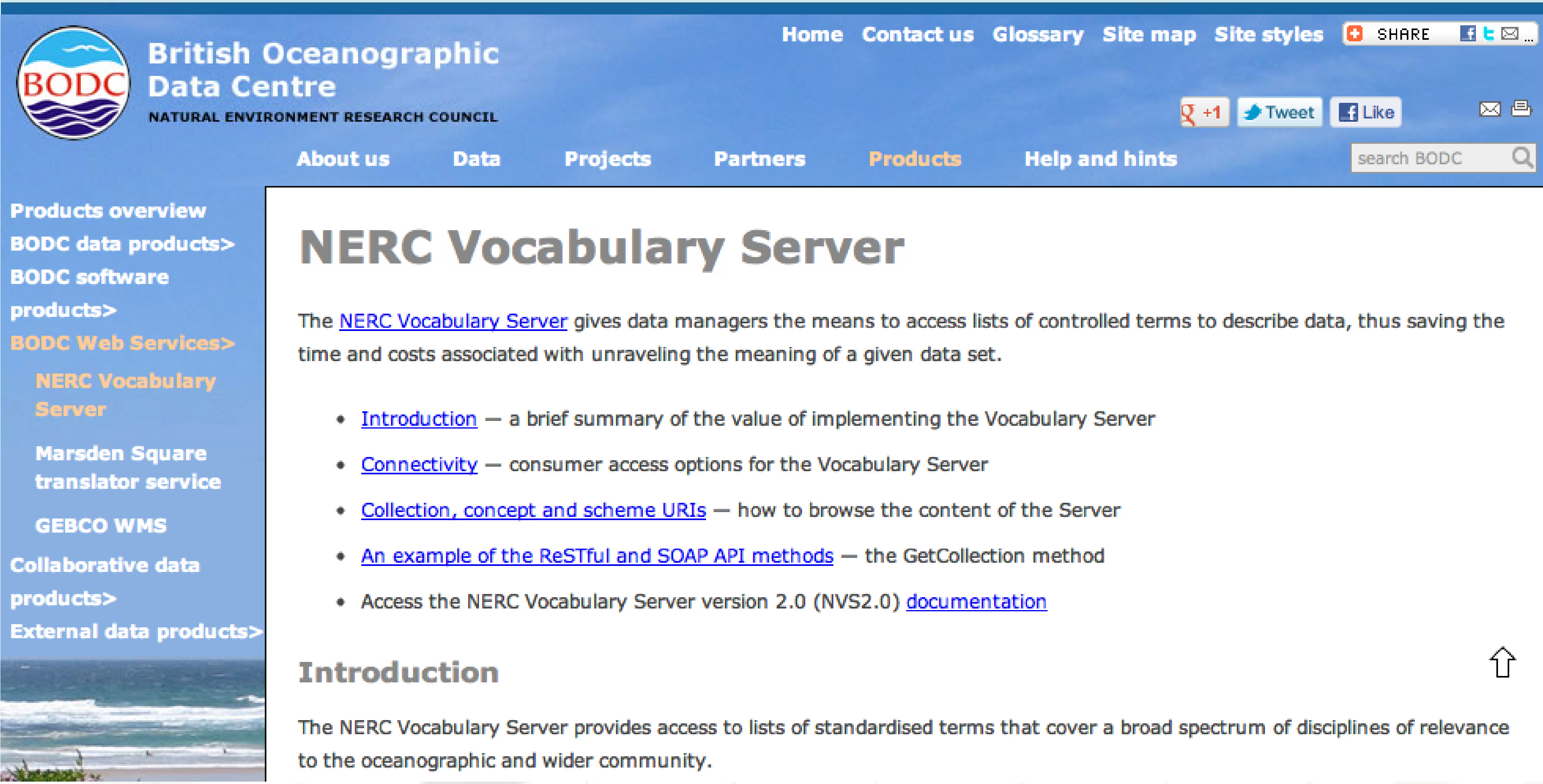
Ontology Services

- SEAVOX
  - http://www.bodc.ac.uk/products/web_services/vocab/
- SEEGRID
  - https://www.seegrid.csiro.au/wiki/Siss/VocabularyService
  - https://www.seegrid.csiro.au/wiki/SISS4BoM/SISSVoc
  - https://www.seegrid.csiro.au/wiki/SISS4BoM/VocabConfiguration
- GeoNetwork
  - http://trac.osgeo.org/geonetwork/wiki/proposals/DCATandRDFServices

Ontology Services



SEAVOX



British Oceanographic Data Centre
NATURAL ENVIRONMENT RESEARCH COUNCIL

Home | Contact us | Glossary | Site map | Site styles | SHARE

+1 | Tweet | Like

About us | Data | Projects | Partners | Products | Help and hints | search BODC

- Products overview
- BODC data products>
- BODC software products>
- BODC Web Services>
  - NERC Vocabulary Server
  - Marsden Square translator service
  - GEBCO WMS
- Collaborative data products>
- External data products>

NERC Vocabulary Server

The NERC Vocabulary Server gives data managers the means to access lists of controlled terms to describe data, thus saving the time and costs associated with unraveling the meaning of a given data set.

- Introduction — a brief summary of the value of implementing the Vocabulary Server
- Connectivity — consumer access options for the Vocabulary Server
- Collection, concept and scheme URIs — how to browse the content of the Server
- An example of the ReSTful and SOAP API methods — the GetCollection method
- Access the NERC Vocabulary Server version 2.0 (NVS2.0) documentation

Introduction

The NERC Vocabulary Server provides access to lists of standardised terms that cover a broad spectrum of disciplines of relevance to the oceanographic and wider community.

SEEGRID

Solid Earth and Environment GRID

SEE GRID community website

Solid Earth and Environment GRID

CSIRO

Jump | Search

SISS4BoM

Log In or Register

- Create personal sidebar

SISS4BoM Web
Create New Topic
Index
Search
Changes
Notifications
RSS Feed
Statistics
Preferences

Webs
ASRDC
AUKEGGS
AnalyticalGeoscience
AppSchemas
CGIModel
Compsrvices
GPML
Gazetteer
Geohazards
GeothermalData
Infosrvices
Know
Main
Marineweb
NRInfo
NeCTARProjects
SCENZGrid
SISS4BoM
Sandbox
Siss
System

You are here: SEEGrid > SISS4BoM Web > SISSVoc > VocabConfiguration (10 Aug 2012, FlorenceTan)

Edit | Attach

Vocab Configuration

Introduction

In this activity, SISS team will create configuration and associated documentation required for SISSVoc 3.x deployment and work with BOM to deploy/understand the service and deploy into Production.

Workflow

- Step 1 : Prepare and Pre-Condition the Vocabulary as an RDF/XML Document
- Step 2 : Publish the Vocabulary as a Single Document
- Step 3 : Create SPARQL Endpoint thru Loading the Vocabulary in OpenRDF from the Ontology URI
- Step 4 : Prepare ELDA Configuration and Deploy to SISSvoc Service
- Setp 5 : Configure the Concept URI Server to Get Concept Descriptions From SISSvoc

Step 1 : Prepare and Pre-Condition the Vocabulary as an RDF/XML Document

- The role of a vocabulary owner is to formalize and precondition the vocabs. BOM as vocabulary owner has identified to

BGS

Our data
OpenGeoscience
Maps & spatial data
Data collections
- Borehole materials
- Discovery metadata
- Earth Science Academic Archive
- GeoIndex
- Groundwater level information
- Lexicon of rock units
- Linked Data
- PalaeoSaurus
- Rock classification
- Rock collections
- Strategic Environmental Assessment
- Taxonomy Online
- Vocabularies

Digital images &

Vocabulary web service — version 1.0

This document describes version 1.0 of the BGS Vocabulary Service.

Contents

- Usage
- Resources
- Supported media types
- Status codes
- Clients

Usage

The vocabulary service has been implemented as a RESTful web service using the Restlet framework.

Information is exposed through a series of resources, where each resource implements the standard HTTP interface and is identified by a URL. Each resource provides representations of information for a number of supported media types.

The vocabulary service supports content negotiation to provide the most appropriate representation for a particular client. For example to request a JSON representation of a resource, a client should use the following HTTP header in its request:

```
Accept: application/json
```

Not all clients will be able to add or modify HTTP headers, therefore the vocabulary service supports other ways to specify the required representation format. Each media type has been mapped to a file extension. Each resource supports a media parameter in the query string which may be used to tunnel the required media type. A JSON representation would be returned by the following:

```
/vocabularies?media=json
```

Additionally like in Ruby on Rails, the file extension may be appended to the URL:

```
/vocabularies.json
```

Resources

All resources are relative to **http://webservices.bgs.ac.uk/data/services/vocabulary/1.0**

See also
- Web services
- Lexicon of rock units
- Rock classification scheme
- Vocabularies

External links
- OneGeology

Share this page
Tweet Share

News and apps

Tell us what you think

Don’t forget - the metadata

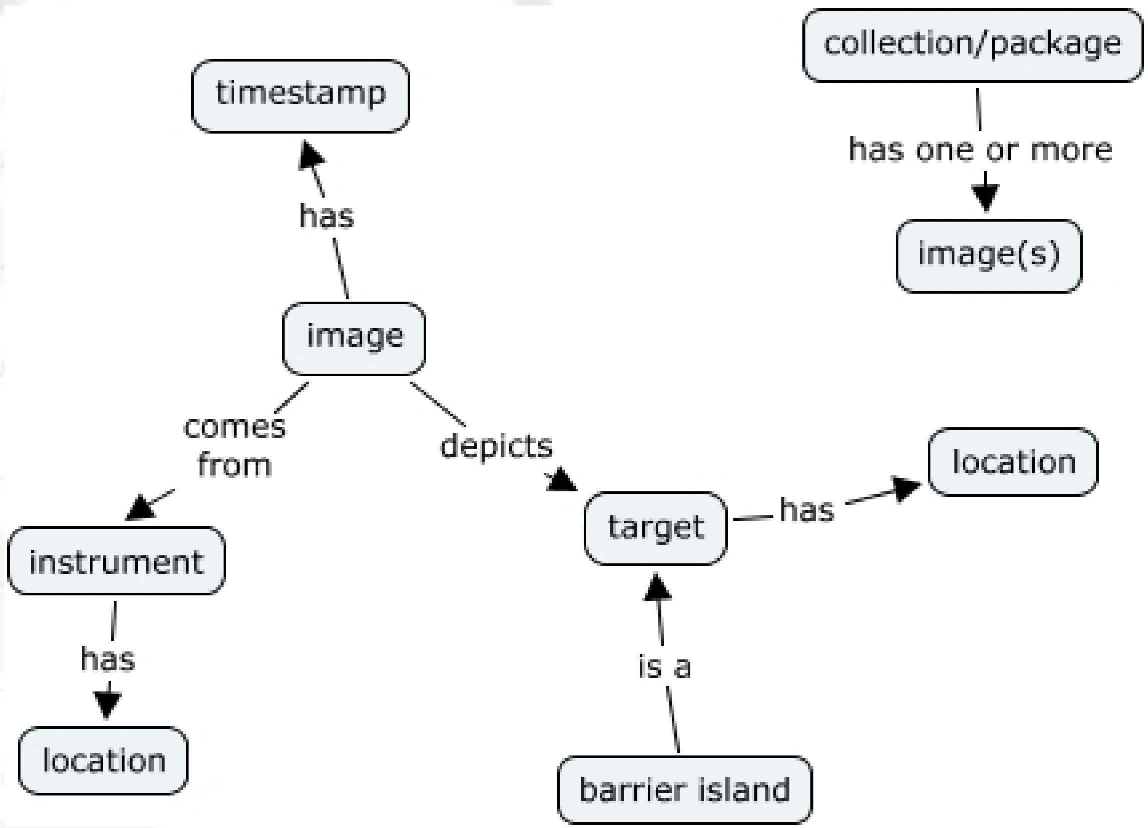
- Metadata (maybe data too) is “materialized” into instances of the ontologies and accessed via services
- I.e. linked in a machine processable (e.g. queryable) way
- And… linked among vocabularies (more general than “mapping”)
- For discovery, inventory, access (and use…) but only what’s needed…

Status/ schedule

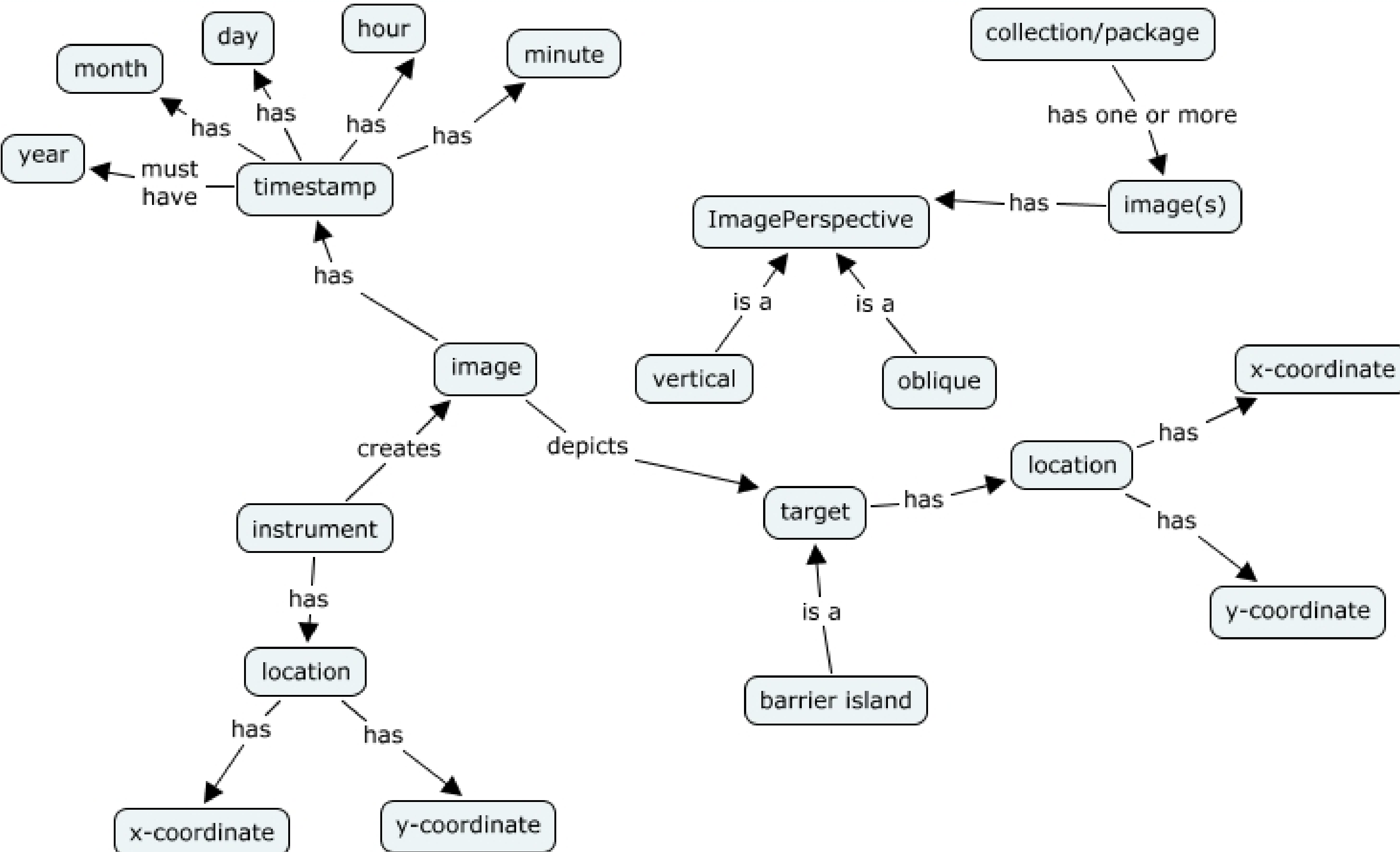
- First iteration of use cases are being analyzed and modeled
- Early fall – first models and ontologies (expect lots of re-use, minimal development), design of vocab service (re-use), prototype application to use it
- Then – evaluation with GeoPortal, etc..

Questions?

- Thanks…
- Peter Fox pfox@cs.rpi.edu @taswegian



collection/package
has one or more
image(s)
timestamp
has
image
comes from
depicts
instrument
has
location
target
has
location
is a
barrier island



collection/package
has one or more
image(s)
has
ImagePerspective
is a
vertical
is a
oblique
month
day
hour
minute
year
must have
timestamp
has
has
has
has
has
image
creates
instrument
has
location
has
x-coordinate
has
y-coordinate
depicts
target
has
location
has
x-coordinate
has
y-coordinate
is a
barrier island