Winter Meeting 2013

Lessons learned in deploying a cloud-based knowledge platform for the ESIP Federation

Line C. Pouchard,* Andrew DePriest ,** Michael N. Huhns**

* Oak Ridge National Laboratory, ** University of South Carolina

A very diverse data and metadata ecosystem

Objectives

- Simulations, models, experiments, remote sensing, GIS, molecular and -omic databases, publications
- Numerous metadata schemas
- Workflows, scenario development, data and process re-use, provenance
- Complex systems of systems, networks of projects, repositories, archives, publishers
- Enable earth and environmental science ontologies to be accessible from the Web
- Host ontology services in a scalable public cloud
- Incorporate tools for managing, accessing, searching, browsing, and disseminating ontologies

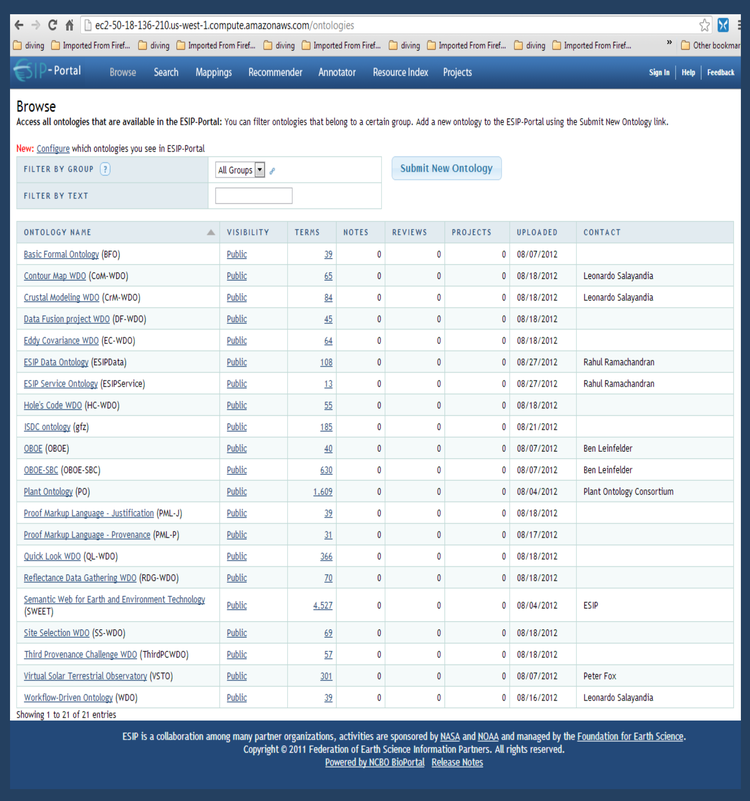
Approach

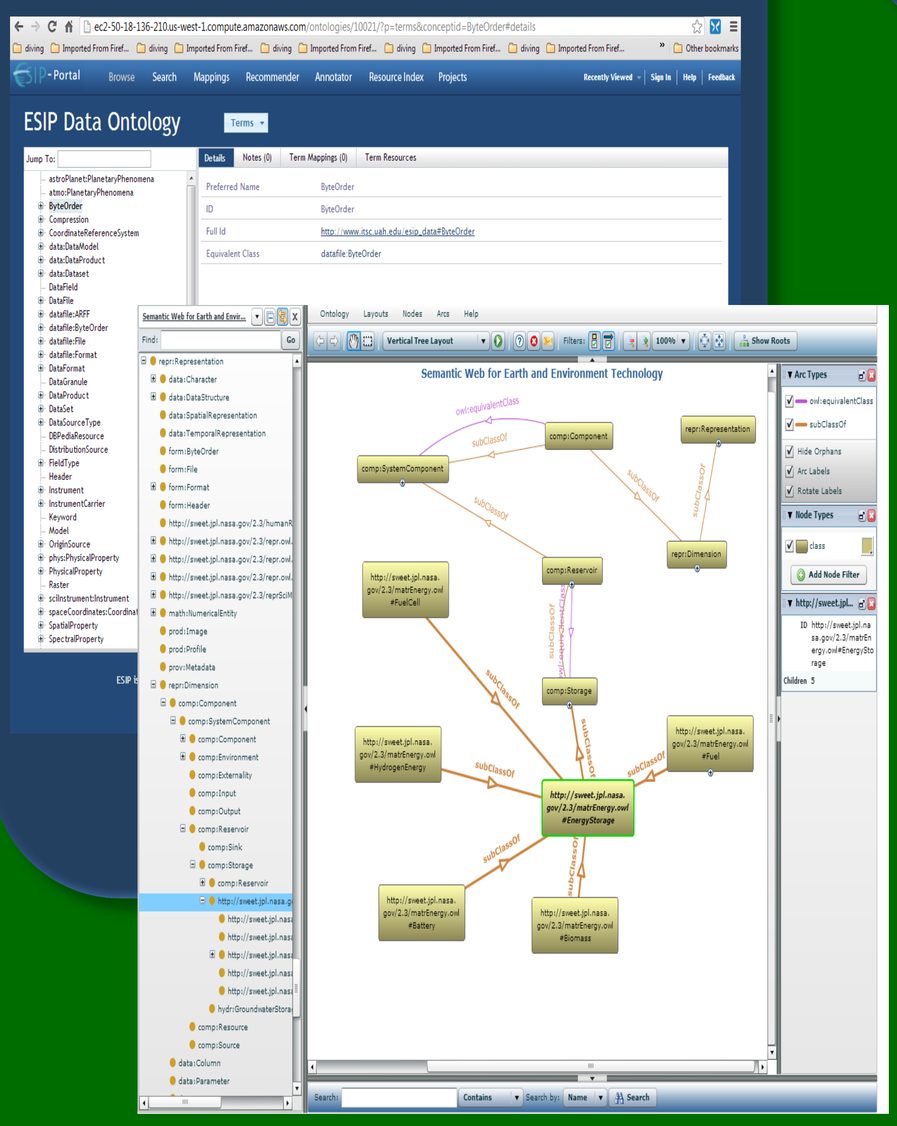
- Investigate and set up a private cloud computing environment at USC for serving ESIP ontologies
- Investigate and set up an Amazon AWS point of service for serving ESIP ontologies
- Obtain and install the BioPortal image
- Produce ESIP-related descriptions for re-branding of BioPortal into an ESIP portal
- Produce outreach material to publicize and assist users with the ESIP-Portal

Advantages

- A repository with tools where ESIP members can store, visualize, share, and map their ontologies
- Ontology versioning managed within the portal
- Community sourcing of ontology maintenance
- Low-budget solution

Implemented Portal





Challenges

- Maintenance
- Governance
- Integration into ESIP infrastructure
- Advertisement of portal contents
- Reconciliation of conflicts among ontologies
- Federation of ontologies

What worked

- Very small budget
- Volunteer effort for the mentors
- Work with undergraduates
- Pre-existing relationships
- Platforms designed for interoperability and re-use
- Open source APIs
- Straightforward deployment in the AWS Elastic Cloud (EC2)
- Required creation of custom image

Collaborations

- ESIP Federation Semantic Web Cluster
- ESIP Federation Products and Services
- National Center for Biomedical Ontologies
- University of South Carolina – Center for IT
- ORNL – Scientific Data Group

What did not work

- Custom image doesn't contain all features
- Lack of continuity
- Lack of incentives for mentors
- Need for better evaluation of ontologies

This work has been in part performed at Oak Ridge National Laboratory,
Managed by UT Battelle, LLC under Contract No.
De-AC05-00OR22725 for the U.S. Department of Energy