The eScience Revolution: Creating Semantic Web Platforms for Massive Scientific Collaboration (tw.rpi.edu/portal/SESF) Deborah McGuinness (dm@cs.rpi.edu), Peter Fox (pfox@cs.rpi.edu), Rensselaer Polytechnic Institute, Troy, NY 12180. Funded by NSF/FOCI award OCI-0943761.

Abstract

Vision: this project seeks to create virtual collaboratories designed to break science data and information out of the laboratory and place it in the hands of the people (http://bit.ly/7HMC8C).

Goal: is to hasten scientific discovery and innovation by enabling rapid and easy collaboration between scientists, educators, students, policy makers, and "citizen scientists" around the world via the Web.

Uses: real life use cases (e.g. Restoration and management of large aquatic ecosystems) to advance innovation and application of semantic web, web and web technologies to enable the compilation and sharing of scientific data on an unprecedented scale.

Approach: Semantic technologies are used to lower the barrier of entry to do science and accelerate the growth of community knowledge, bridging gaps between questions that someone wants to ask in their limited scientific vocabulary and the extreme complexity of the underlying data.

Provides: a toolkit for scientists, policy professionals, and educators allowing them access to data from a variety of sources and, importantly, outside of their direct area of expertise but customized with vocabularies suitable for their use.

Behind: the interfaces, the toolkit uses semantic ontologies to customize Web sites, services and programming interfaces.

Look: Web sites will be familiar, understandable, and navigable to end users depending on the level and type of expertise. The user needs only to type a question, and it may be answered using data input customization and type of expertise. The user needs only to type a question, and it may be answered using data input customization and type of expertise. The user needs only to type a question, and it may be answered using data input customization and type of expertise. The user needs only to type a question, and it may be answered using data input customization and type of expertise.

Funded: by the American Recovery and Reinvestment Act via NSF/FOCI Strategic Technologies for CyberInfrastructure (STCI)

Semantic data framework architecture serving multiple communities

Prior to 2005, we built systems, now we build frameworks

Rough definitions

Systems have very well-defined 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, exit, and use points. A user often does not know when they are using one. Extension points are part of the design.

Framework v. Systems

Non-specialist Use-Case(s): e.g. Access to Science Data from Virtual Observatories

The Virtual Solar-Terrestrial Observatory (VSTO) utilizes leading edge virtual observatory capabilities among three sub-disciplines: solar radiation, volcanic outgassing and atmospheric structure using extensions to existing modular ontologies and used the VSTO data framework, while adding smart faceted search and semantic data registration tools (NASA; 3 years).

Semantic Provenance Capture in Data Ingest Systems (SPDCIS) has added explanation provenance capabilities to an observational data ingest pipeline for images of the Sun providing a set of tools to answer diverse end user questions such as: "Why does this image look bad?" (NSF; 2 1/2 years).

Key Technology needs

- The ontology web language - OWL 2.0 and it's profiles, especially OWL 2 RL
- A services language SAWSDL and OWL-S
- Ontology editors – Protégé 4.0 and CMap
- An ontology reasoner – Pellet 2.0, Jena, Jess
- Semantic query – SPARQL – Triple store – Jena, Virtuoso, etc.
- An explanation interlingua - PML - proof markup language
- A suitable explanation infrastructure - Inference Web (http://inference-web.org)
- A collaborative portal infrastructure – Drupal 6/7
- Application-level semantic tools; faceted search, provenance explanation and browse

More Information: http://tw.rpi.edu/portal/SESF