Enriching Ferryboxes on the Semantic Web for a Collaborative Ocean

Feng (Barry) Tao Jon Campbell
National Oceanography Centre, Southampton

Collaborative Ocean is an initiative under Oceans2025 to build a semantic web infrastructure to enable marine scientists to collaboratively publish their resources on the semantic web to assist multi-use of oceanographic data. Intelligent resource discovery, inference and data interoperability are some of the benefits that should flow from adopting a semantic web based approach. The initiative started within the National Oceanography Centre, Southampton. We hope interested oceanographers, marine engineers and data providers can work collaboratively with the semantic web community to help opening up their data through weaving the semantic web.

Smart Sensors – The Vision

•Sensors should be self-describing and able to automatically provide/reuse semantic annotations to identify themselves and be aware of contextual information in their residential sensor network
•Data structures should also be self-describing with machine readable headers defining the data types, formats and metadata
•By following internationally agreed standards, data would be immediately accessible to other organizations around the world
•Time taken to interface new sensors and adapt software to handle new data types is greatly reduced

Example of a Smart Sensor - MBARI PUCK

PUCK helps to automate the configuration process by physically storing information about the instrument with the instrument itself. The information may include metadata descriptions, driver software, GUI software, or any other information deemed relevant by the observing system. See http://www.mbari.org/pw/puck.htm

Who are we collaborating with externally?

Although most effort is currently confined to describing resources within the NOCS, we also work together with Marine Metadata Interoperability (MMI), in particular their device ontology group, for ontology development. These ontologies can be reused in the semantic web infrastructure. We also collaborate with the BODC vocabulary to reuse existing concepts and demonstrate data interoperability.

We are happy to hear more from other organizations having similar interests or wishing to describe/publicize their resources on the semantic web.

Why should different groups contribute to the Collaborative Ocean ?

Data/Process managers – Improved ability to manage and maintain measurement datasets and associated metadata, instruments and equipment.
Oceanographic engineers – Developing smart sensors and smart interfaces for existing sensors makes sensor integration easier and improves management of calibrations and other metadata.
Environment scientists - Improved dataset discovery services according to different search criteria. Better scientific data provenance, i.e., improved linkage between experimental oceanographic data and scientific analysis.
Knowledge engineers - Responsible of elicitation of domain requirements, choosing knowledge engineering methodologies and setting up knowledge management activities within the knowledge life cycle, e.g., ontology engineering, semantic annotation and semantics reuse. Working together with other participants to establish the semantic web infrastructure
Software developers - Technical development of software and integration of different sub-systems to fulfil knowledge management requirements in the domain.

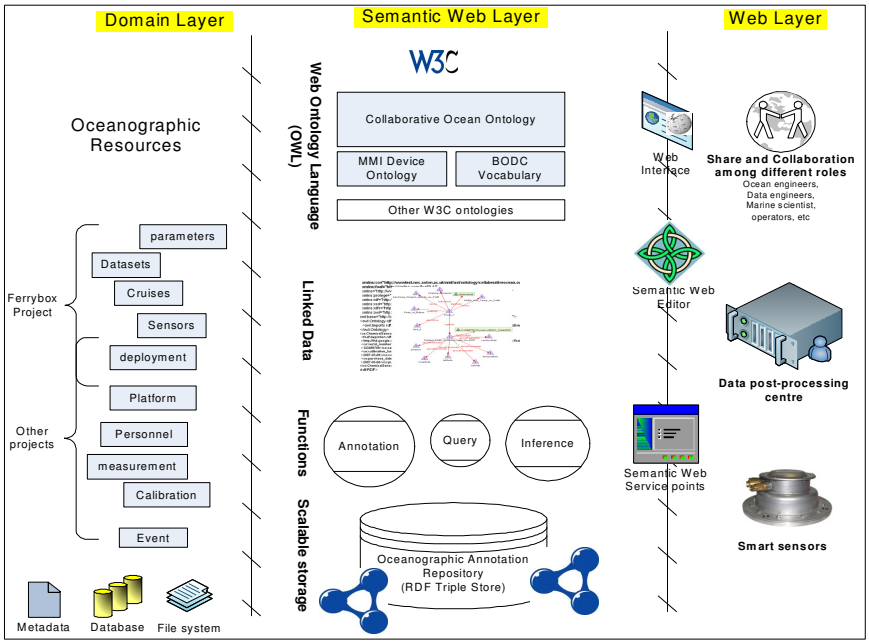
Potential Application to Ferrybox System

• Smart sensors can automatically provide/reuse semantic annotations while logging measurement data
• Improve management of sensors and their calibrations
• Annotate a dataset to link with sensor and other resource
• Discover and plot data from common sensor types
• Facilitate device classification according to functionality
• Plot datasets on the map using geospatial annotations

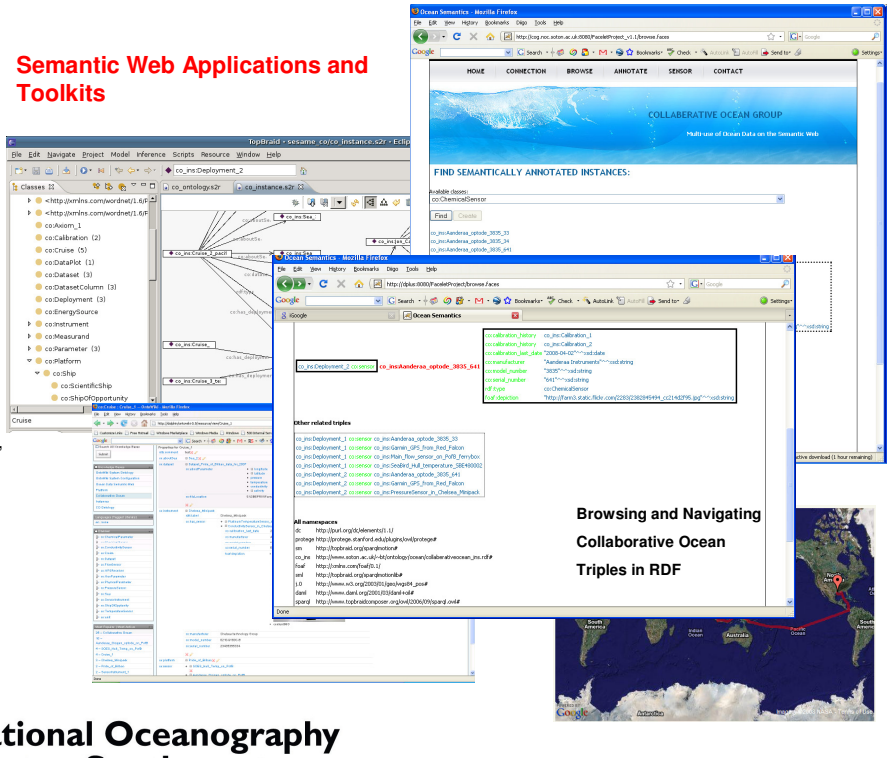
What is the Semantic Web?

Semantic web is a web of linked data that follows W3C semantic web related specifications. In particular, RDF (Resource Description Framework), based on XML syntax, is the backbone to model knowledge as a graph of nodes and linkages with defined namespaces in Ontology. Models represented in the RDF graph can be represented as triples, which improves consistency and machine processing and maximises the efficiency of distributed storage.

System Scenario



Semantic Web Applications and Toolkits



National Oceanography Centre, Southampton
UNIVERSITY OF SOUTHAMPTON AND NATURAL ENVIRONMENT RESEARCH COUNCIL

Contact email: bt@noc.soton.ac.uk