



A semantically rich and standardised approach enhancing discovery of sensor data and metadata

Alexandra Kokkinaki, Justin Buck, and Louise Darroch

The British Oceanographic Data Centre, Liverpool, United Kingdom (alexk@bodc.ac.uk)

The marine environment plays an essential role in the earth's climate. To enhance the ability to monitor the health of this important system, innovative sensors are being produced and combined with state of the art sensor technology. As the number of sensors deployed is continually increasing, it is a challenge for data users to find the data that meet their specific needs. Furthermore, users need to integrate diverse ocean datasets originating from the same or even different systems.

Standards provide a solution to the above mentioned challenges. The Open Geospatial Consortium (OGC) has created Sensor Web Enablement (SWE) standards that enable different sensor networks to establish syntactic interoperability. When combined with widely accepted controlled vocabularies, they become semantically rich and semantic interoperability is achievable. In addition, Linked Data is the recommended best practice for exposing, sharing and connecting information on the Semantic Web using Uniform Resource Identifiers (URIs), Resource Description Framework (RDF) and RDF Query Language (SPARQL).

As part of the EU-funded SenseOCEAN project, the British Oceanographic Data Centre (BODC) is working on the standardisation of sensor metadata enabling 'plug and play' sensor integration. Our approach combines standards, controlled vocabularies and persistent URIs to publish sensor descriptions, their data and associated metadata as 5 star Linked Data and OGC SWE (SensorML, Observations & Measurements) standard. Thus sensors become readily discoverable, accessible and useable via the web. Content and context based searching is also enabled since sensors descriptions are understood by machines. Additionally, sensor data can be combined with other sensor or Linked Data datasets to form knowledge.

This presentation will describe the work done in BODC to achieve syntactic and semantic interoperability in the sensor domain. It will illustrate the reuse and extension of the Semantic Sensor Network (SSN) ontology to Linked Sensor Ontology (LSO) and the steps taken to combine OGC SWE with the Linked Data approach through alignment and embodiment of other ontologies. It will then explain how data and models were annotated with controlled vocabularies to establish unambiguous semantics and interconnect them with data from different sources. Finally, it will introduce the RDF triple store where the sensor descriptions and metadata are stored and can be queried through the standard query language SPARQL.

Providing different flavours of machine readable interpretations of sensors, sensor data and metadata enhances discoverability but most importantly allows seamless aggregation of information from different networks that will finally produce knowledge.