A service for publishing sensors on the web using OGC and W3C standards

Louise Darroch, British Oceanographic Data Centre, National Oceanography Centre (UK), louise.darroch@bodc.ac.uk
Alexandra Kokkinaki, British Oceanographic Data Centre, National Oceanography Centre (UK), alexk@bodc.ac.uk
Justin Buck, British Oceanographic Data Centre, National Oceanography Centre (UK), juck@bodc.ac.uk
Thomas Gardner, British Oceanographic Data Centre, National Oceanography Centre (UK), thogar@bodc.ac.uk
Cieran Brazier, British Oceanographic Data Centre, National Oceanography Centre (UK), cbraz@bodc.ac.uk
Gwenaelle Moncoiffé, British Oceanographic Data Centre, National Oceanography Centre (UK), gmon@bodc.ac.uk

Autonomous ocean observation is massively increasing the number of sensors in the ocean. Advances in this data gathering technology mean that we are generating more data than ever before, presenting challenges in data automation, handling geospatial information and linking observations to sensor metadata, the sensor's context that supports scientists on selecting quality sensors and their data. The Open Geospatial Consortium (OGC) has developed the Sensor Web Enablement (SWE) standards to facilitate integration and interoperability of sensor data and metadata. World Wide Web Consortium (W3C) Semantic Web technologies enable machine comprehensibility promoting sophisticated linking and processing of data published on the web. However, there are some practical difficulties when linking sensor content and context using the above-mentioned standards, because of internal hardware bandwidth restrictions and a requirement to constrain data transmission costs. Furthermore, Observations and Measurements (O&M), one of the core standards in the SWE suite, defines a conceptual schema encoding for observations. However, it was primarily designed for handling observations on a data point-by-point basis, requiring administrators to encode data with different dimensions and formats on a case-by-case basis, thereby reducing data integration and automation.

As part of the EU project, SenseOCEAN, the Natural Environment Research Council's (NERC) British Oceanographic Data Centre (BODC) developed the Marine Linked Systems (fig. 1), a SWE and World Wide Web Consortium (W3C) compliant sensor web publication service. Our approach addresses the practical difficulties with hardware bandwidth and transmission costs by uniquely identifying sensor and platform models and instances through URIs that act as Universally Unique Identifiers (UUIDs), which resolve via content negotiation to either OGC's sensor meta-language, sensorML or W3C's Linked Data enriched with Semantic Sensor Network (SSN) ontology. Sensor and platform model URIs and descriptions are created and hosted by the service. Sensor and platform instance URIs are dynamically created prior to and during sensor deployments, by the sensor owner, and are associated with the relevant model URI through an updatable web form, the Sensor Instance Form. Association between platform and sensor URIs is also performed by the end user, representing deployment. When sensors transmit their content, they include their unique URI to refer to their content (fig. 1).

The service exposes multi-dimensional and formatted sensor observations using OGC's O&M. However, the model used is based on the 'out-of-band' principle, which points to external data resources including Digital Object Identifiers (DOIs) whilst still enabling the response model for the

OGC Sensor Observation Service (SOS), in this instance a 52north SOS. Thus, different data types can be added to the system with minimal effort.

In addition, semantic interoperability is enhanced by constraining values using internationally established standardised lists of terms (controlled vocabularies), such as the BODC Parameter Usage Vocabulary (P01), that are published on the NERC Vocabulary Server (NVS). The NVS is a publicly available service for the marine community based on the W3C Simple Knowledge Organization System (SKOS) where each term has a unique URI that is resolvable through a RESTful interface to either HTML or RDF documents through content negotiation.

The use of URIs and the Sensor Instance Form offers both practical and economical benefits to the implementation of SWE and Linked Data standards in near real time systems. Data can be linked to metadata dynamically in-situ while saving on the costs associated to the transmission of long metadata descriptions. The transmission of short URIs also enables the implementation of standards on systems where it is impractical, such as legacy hardware. Exposing sensor observations using O&M based on the 'out-of-band' principle provides a way to easily integrate data into SWE compliant environmental networks and may be particularly cost-effective for sensor systems generating large data volumes where it may not be practical to encode each data point. Using established and web-resolvable controlled vocabularies helps to harmonise sensor data and information across these networks.

The development of this service is supported by the EU projects, SenseOCEAN, AtlantOS, EMSODev, ENVRIplus, BRIDGES and EMODNet.

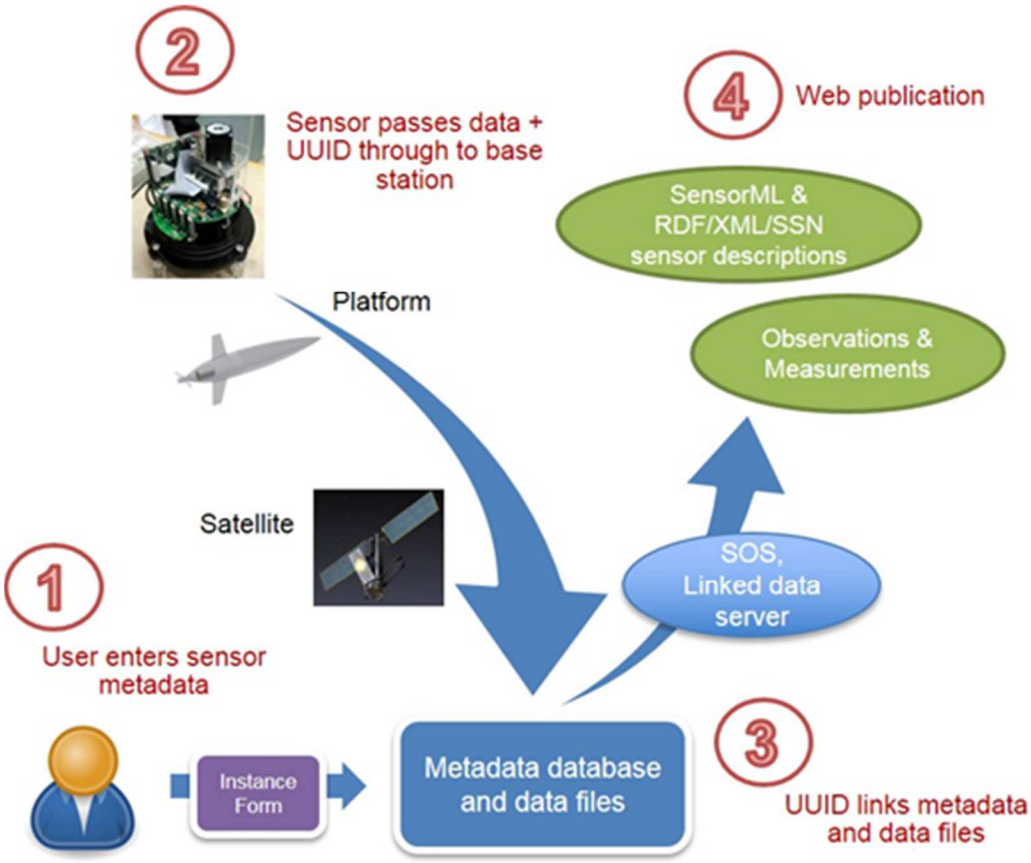

Figure 1: The BODC Marine Linked Systems