We have “born digital” – now what about “born semantic”?

Adam Leadbetter & Janet Fredericks
Overview

- Motivation
- Qartod to OGC
- Lake Ellsworth
- SenseOCEAN and future work
- Conclusions
Motivation

- Bainbridge (2012) @ EGU
  - Cloud-based service-orientated data system for ocean observational data - an example from the coral reef community
- Sparked the question
  - If we have so many observations, can we semantically annotate them from collection?
Motivation

- Groundwork
- Prototyping SeaBird CTD XML with RDFa

```xml
<Sensors xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:muo="http://purl.oclc.org/NET/muo/muo#"
    xmlns="http://www.w3.org/1999/xhtml"
    count="11">
    <sensor Channel="1"
        about="#channel1"
        rel="rdfs:isDefinedBy"
        href="http://vocab.nerc.ac.uk/collection/P01/current/TEMPCU01/">
    <!-- Frequency 0, Temperature -->
    <Units rel="muo:measuredIn"
        href="http://vocab.nerc.ac.uk/collection/P06/current/UPAA/" />
    <TemperatureSensor SensorID="55"
        about="#sensor55"
        rel="rdfs:isDefinedBy"
        href="http://vocab.nerc.ac.uk/collection/L22/current/TOOL0214/"/>
```
Motivation

- Groundwork
  - Prototyping SeaBird CTD XML with RDFa

- Synergy
  - Back in the office - contacting manufacturers
  - Work has already taken place!

- Common goal
  - Dynamic, integrated data quality assessment
Community-based Development

What information is needed to assess quality of data and how do we implement it into an Sensor Observation Services (SOS)?

- Domain Experts
- Sensor Mfgrs
- Operators
- IT Specialists
- Content Specifications/SWE Implementation Model (Can it be an OGC Best-Practice?)
<sml:output name="swell">
<swe:Quantity definition="http://mmisw.org/ont/mvco/properties/swell">
<swe:uom code="cm"/>
</swe:Quantity>
</sml:output>
Lake Ellsworth

- Collaboration with sensor laboratory
  - Mapping from file header to SKOS concepts on NVS
  - Inserted into instrument firmware
  - Sadly project currently on hold
Lake Ellsworth

Pressure (decibars)
- PPSCZZ01 if not corrected to read zero at lake surface, otherwise PRESPR01.

Temperature (°C)
- TEMPST01 if temperature scale unknown.
  (TEMPS601 if IPTS-68 / TEMPS901 if ITS-90)

Conductivity (mS/cm)
- CNDCST01 (we use S/m as the units but that's not an issue).

Salinity (PSU)
- PSALCU01

Fresh water conductivity
- No code - we can create one if needed.

Computed conductivity @ 20°C
- No code - we can create one if needed.

Oxygen saturation (%)
- OXYSZZ01

Oxygen concentration (ppm) and conversion to µmol based on pH
- DOXMZZXX (assuming mg/kg is equivalent to ppm known weight of oxygen for BODC units)

Eh (mV)
- PHXXPR01
- No code - we can create one if needed (I'm assuming this is reduction potential)
SenseOCEAN
SenseOCEAN

- Collaboration with sensor laboratory / commercial manufacturers
  - Assign URLs to SKOS concepts from collection
  - Semantically annotated SensorML
- Issues:
  - Data logger / communications bandwidth
  - Easy on a base station / mooring – not so on an AUV
  - Communicate changes in state only
SeaDataNet

```xml
<swe:Vector definition="gpsLocation" referenceFrame="http://www.example.com">
  <swe:coordinate name="Latitude">
    <swe:Quantity definition="ALATGP01" axisID="latitude">
      <swe:uom code="deg"/>
    </swe:Quantity>
  </swe:coordinate>

  <swe:coordinate name="Longitude">
    <swe:Quantity definition="ALONGP01" axisID="longitude">
      <swe:uom code="deg"/>
    </swe:Quantity>
  </swe:coordinate>
</swe:Vector>
```
Conclusions

- Benefits
  - Data are described well enough for assessment of data quality for specified use and for a repurposed application from the point of collection
  - Hook into ontology development
    - O&M, PROV, Ocean Data Ontology
Conclusions
Conclusions

- **Benefits**
  - Data are described well enough for assessment of data for specified use and for a repurposed application from collection
  - Hook into ontology development
    - O&M, PROV, Ocean Data Ontology
  - Moving into the realm of (near) real-time Linked Data
Conclusions
Conclusions

- Two approaches
  - Community building of specific semantics
  - Or reuse of existing resources
- Common ground
  - Embed semantics in OGC SWE documents
  - Not necessarily communicated across the full network
  - Mapped resources
Conclusions
Conclusions

- Two approaches
  - Community building of specific semantics
  - Or reuse of existing resources
- Common ground
  - Embed semantics in OGC SWE documents
  - Not necessarily communicated across the full network
  - Mapped resources
- Take it to your communities!

alead@bodc.ac.uk / jfredericks@whoi.edu