Observations to Information

2013 Leptoukh Lecture

Simon Cox  |  Research Scientist  |  Environmental Information Systems
12 December 2013
Standards for observation data

• Motivation and development

• Earth-science and environmental applications

• Renewal
Origins
A bit of history

- 1993-2000 AGCRC
  - Web-mapping
  - Reporting research online

- 1995
A bit of history

- 1993-2000 AGCRC
  - Web-mapping
  - Reporting research online

- 1999-2004 XMML, ADX
  - Exploration data standards
A bit of history

• 1993-2000 AGCRC
  • Web-mapping
  • Reporting research online

• 1999-2004 XMML, ADX
  • Exploration data standards
A bit of history

• 1993-2000 AGCRC
  • Web-mapping
  • Reporting research online

• 1999-2004 XMML, ADX
  - 2010 AuScope
  • Exploration data standards
A bit of history

- 1993-2000 AGCRC
  - Web-mapping
  - Reporting research online

- 1999-2004 XMML, ADX
  - 2010 AuScope
  - Exploration data standards

- 2002-2005 OGC SWE
  - Sensors anywhere
A bit of history

• 1993-2000 AGCRC
  • Web-mapping
  • Reporting research online

• 1999-2004 XMML, ADX - 2010 AuScope
  • Exploration data standards

• 2002-2005 OGC SWE
  • Sensors anywhere

• 2005-2013 Fluid Earth
  • Water informatics
A bit of history

- 1993-2000 AGCRC
  - Web-mapping
  - Reporting research online

- 1999-2004 XMML, ADX
  - 2010 AuScope
  - Exploration data standards

- 2002-2005 OGC SWE
  - Sensors anywhere

- 2005-2013 WIRADA
  - Water informatics
Motivation for a standard

All of society’s grand challenges require data to be shared and integrated across cultures, scales and technologies.
Motivation for a standard

- Integrated analysis and modelling

- Discovery & data integration a significant challenge

- Standard vocabulary
  
  **Many private contracts one public agreement!**
SWE/O&M
An Observation is an action whose result is an estimate of the value of some property of the feature-of-interest, obtained using a specified procedure.
Scope

- In situ observations
- Remote sensing
- Ex-situ observations
- Numerical models/simulations
- Forecasts

Any action whose result is an *estimate* of a property value
Sampling features

Harmonized with CSML, NCAR

Cox, OGC Abstract Specification – Topic 20: Observations and Measurements 2.0
ISO 19156:2011 Geographic Information – Observations and measurements
OGC Sensor Web Enablement

- SensorML
- O&M

- Sensor Observation Service
- Sensor Planning Service
Sensor Observation Service

getObservation
getResult
describeSensor
getFeatureOfInterest
Observations & Measurements

<table>
<thead>
<tr>
<th>procedure</th>
<th>result</th>
</tr>
</thead>
<tbody>
<tr>
<td>observed property</td>
<td>feature of interest</td>
</tr>
<tr>
<td>result time</td>
<td>phenomenon time</td>
</tr>
<tr>
<td>valid time</td>
<td></td>
</tr>
</tbody>
</table>
Integrated into ISO 19100 framework

ISO dependencies

- gf - Feature ISO 19109
- cv - Coverage (fields) ISO 19123
- md - Metadata ISO 19115
- gm - Geometry ISO 19107
- tm - Temporal ISO 19108
- basic - Datatypes ISO 19103
Domain specialization

- **feature of interest**
  - Feature-type catalogue
  - Feature service

- **procedure**
  - Register of sensors, processes & algorithms

- **result format:**
  - GML, SWE, netCDF, JSON, SQLite ...

---

- **OM_Observation**
  - phenomenonTime
  - resultTime
  - validTime [0..1]
  - resultQuality [0..*]
  - parameter [0..*]

- **GFI_PropertyType**

- **OM_Process**
  - procedure

- **GFI_Feature**
  - featureOfInterest
  - propertyValueProvider
  - observedProperty

- **GFI_DomainFeature**
Linked vocabularies can be shared and re-used
Applications
WaterML2 specializes O&M

- Result is a **time-series**
- Observed-property relates to **water**

- Key predecessors:
  - CUAHSI WaterML
  - WDTF
WaterML-WQ constrains O&M and WaterML

- Subject is a groundwater or geofabric feature
- Observed property is water-quality property
- Units of measure must match
Observables vocabulary

- **nitrogen**
  - **dissolved nitrogen concentration**
    - **elemental nitrogen (CHEBI:33267)**
      - **MolePercent**
        - **MilliGramsPerLitre**

**QUDT**

**ChEBI**
ANZSoilML extends O&M Sampling Features

... also GeoSciML
Motivation

From instrument to User

Existing Activity

- Proprietary: No IP, No
- Permanent/No: IP/No
- Broadband/No
- Fully IP & Broadband

Diagram:

Data Service

Final User

Instrument

Satellite link

Vessels Acq. System

INSTRUMENTS

Buys/Observatories Acq. System

Platforms Operator HQ

Data Centres

SeaDataNet Portal

SeaDataNet
Existing Activity

SWE Adoption

Seismic Profile O&M

Diviacco, P. et al. 2011 Marine Seismic Metadata for an Integrated European Scale Data Infrastructure
Integration with existing SeaDataNet Standards
EU Air Quality Directive

L 335/86

Official Journal of the European Union
17.12.2011

COMMISSION IMPLEMENTING DECISION
of 12 December 2011

<table>
<thead>
<tr>
<th>IPR Nr</th>
<th>Reference</th>
<th>Element</th>
<th>Specification</th>
<th>Description</th>
<th>Requirement</th>
<th>Cardinality</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.5</td>
<td>The dataset describing Fixed and Indicative Measurements comprises meta-information on three hierarchical levels: Measurement Configuration, Station, Network</td>
<td>M</td>
<td>1. *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.5.1</td>
<td>Information concerning measurement configuration by pollutant (or group of pollutants) to be provided for each measurement configuration for which data is to be reported. According to art. 2 (5), a 'Measurement Configuration' means the technical facility used for the measurement of one pollutant or one of its compounds at a specific location. For a pollutant, more than one Measurement Configuration, at a single location, be in operation.</td>
<td>M</td>
<td>1. *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. D.5.1.1</td>
<td>Measurement configuration ID</td>
<td>Identifier</td>
<td>Unique identifier for the measurement configuration. In case of existing measurement configuration, the «location» should be the Extending code. In case of new measurement configuration the location is to be defined by the responsible body, comprising station code, pollutant «Atmos» code and a number.</td>
<td>M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2. D.5.1.2</td>
<td>Station ID</td>
<td>Link</td>
<td>Station ID from existing codes that identify stations for which monitoring is provided. In case of D.5.1.1, the station ID is also the measurement configuration ID.</td>
<td>M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3. D.5.1.3</td>
<td>Network ID</td>
<td>Link</td>
<td>Network which manages the sampling point.</td>
<td>M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D.5.4</td>
<td>Measurement meta references</td>
<td>M</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. D.5.4.1</td>
<td>Measurement start date</td>
<td>Metadata</td>
<td>Year of the measurement configuration (measurement of a pollutant at a monitoring station).</td>
<td>M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2. D.5.4.3</td>
<td>Measurement end date</td>
<td>Metadata</td>
<td>Year of the measurement configuration (measurement of a pollutant at a monitoring station).</td>
<td>M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D.5.5</td>
<td>Emissions with predominant influence. For each measurement configuration, the type of source is to be given which is responsible for the largest relative contribution to the observed concentration. Describe emissions influencing the measurement station for the specific pollutant. This information is relevant for the interpretation of the measured data and the assessment of the representativeness of the site.</td>
<td>M</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. D.5.5.1</td>
<td>Classification of stations in relation to major emission sources relevant for the measurement configuration</td>
<td>multi</td>
<td>Use pollutant</td>
<td>Classification of stations in relation to dominant-emission sources defined in the Exchange of Information (97/101/EC).</td>
<td>M</td>
<td>1</td>
</tr>
<tr>
<td>23. D.5.5.2</td>
<td>Main emission source</td>
<td>multi</td>
<td>Use pollutant</td>
<td>The main emission source(s) (D.5.5.2) can be selected from the Codified Main Emission Sources. In addition to the CRF emission database, two 'source' types have been introduced which represent contributions not originating from identifiable sources: ‘Secondary’ and ‘Long-range’.</td>
<td>C</td>
<td>0.1</td>
</tr>
</tbody>
</table>
D2.9 Draft Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development

Title
D2.9 Draft Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development

Creator
INSPIRE Cross Thematic Working Group on Observations & Measurements

Date
2013-02-22

Subject
Use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE

Publisher
INSPIRE Cross Thematic Working Group on Observations & Measurements

Type
Text

Description
This document describes the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE data specification development.

This version (version 2, release candidate 3) reflects the content of the draft amendment to Commission Regulation (EU) No 1089/2010 for the Annex II/III spatial data themes as submitted to the INSPIRE Committee.
ISO19156 Observations and measurements:
also suitable for numerical simulations – including forecasts

forecast : OM Observation

- parameter.name = “analysisTime”
- parameter.value = 2010-05-06T00:00Z
- phenomenonTime.begin = 2010-05-06T00:00Z
- phenomenonTime.end = 2010-05-09T12:00Z
- resultTime = 2010-05-06T04:30Z
- validTime [optional – not specified]
- resultQuality [optional – not specified]
Common constraints applicable to all WMO METCE Observation types

All specialisations of OM_Observation defined in WMO METCE require:

- association role ‘featureOfInterest’ shall be of type SF_SpatialSamplingFeature
- association role ‘procedure’ shall be of type Process (from WMO METCE)
In meteorology, we define a sampling regime that enables us to observe, measure or simulate the real-world. Sampling Features (from ISO 19156 ‘Observations and measurements’) provide a way to characterise this sampling regime and the relationship to the real-world.
NEII components

Metadata Catalog

Observing Methods Register

Vocabulary Service

Data Services

Monitoring Sites Register
Integration through linking table

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Site</th>
<th>Geography</th>
<th>Parameter</th>
<th>Instrument / Method</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Observation

- feature of interest
- observed property
- procedure
- result
IGSN Sample Metadata

- IGSN
- ResourceURI
- Registrant ID
- timeStamp
- status

ISO & OGC O&M compliant

Allocating Agent Catalog

<table>
<thead>
<tr>
<th>IGSN:</th>
<th>CMY00007W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Name:</td>
<td>TN182_47_002</td>
</tr>
<tr>
<td>Other Name(s):</td>
<td>Sample Type: Individual Sample</td>
</tr>
<tr>
<td>Parent IGSN:</td>
<td>CMY00001B</td>
</tr>
</tbody>
</table>

Description

- Material: Rock
- Classification: Igneous>Plutonic>Mafic
- Field Name: gabbro, hornblende gabbro
- Description: mafic plutonic rock
- Age (min): 2 Ma
- Age (max): 50 Ma
- Collection Method: Dredging
- Collection Method Description: Not Provided
- Size: 0.5 kg
- Geological Age: Cenozoic
- Geological Unit: Not Provided
- Comment: Not Provided
- Purpose: Not Provided

Geolocation

- Latitude: 51.6298
- Longitude: 179.8611
- Elevation: -1460 m
- Physiographic Feature: fault scarp
- Name Of Physiographic Feature: Sunday Basin
- Location Description: north wall of Sunday Basin depression

AGU Fall 2013 | IN42A-01 | Cox | Leptoukh
IGSN Sample in OMXML

<spec:SF_Specimen gml:id="abc123">
  <gml:identifier codeSpace="http://igsn.org">IGSN.SIOabc123</gml:identifier>
  <sa:sampledFeature xlink:href="http://example.org/midAtlanticRidge"/>
  <sa:relatedSamplingFeature>
    <sa:SamplingFeatureComplex>
      <sa:role xlink:href="http://example.org/parent"/>
      <sa:relatedSamplingFeature xlink:href="http://handle.net/10273/IGSN.SIOxyz456"/>
    </sa:SamplingFeatureComplex>
  </sa:relatedSamplingFeature>
  <spec:materialClass xlink:href="http://example.org/rock"/>
  <spec:samplingTime>
    <gml:TimeInstant gml:id="tim123">
      <gml:timePosition>2009-06-30T08:00:00.00-02:00</gml:timePosition>
    </gml:TimeInstant>
  </spec:samplingTime>
  <spec:samplingMethod xlink:href="http://example.org/ghostbuster"/>
  <spec:samplingLocation>
    <gml:Point gml:id="loc123" srsName="http://www.opengis.net/def/crs/EPSG/0/4326">
      <gml:pos>45.12 -81.78</gml:pos>
    </gml:Point>
  </spec:samplingLocation>
  <spec:size uom="kg">0.46</spec:size>
  <spec:specimenType xlink:href="http://example.org/splitCore"/>
</spec:SF_Specimen>
IGSN Sample in RDF

```xml
<http://handle.net/10273/IGSN.SIOabc123>
  a sam:Specimen ;
  rdfs:label "SIO specimen abc123"^^xsd:string ;
  sam:currentLocation <http://example.org/Various/Warehouse3/shelf9/box67> ;
  sam:materialClass p1:rock ;
  sam:preparationStep
    [ sam:processOperator p1:JohnDoe ;
      sam:processingDetails
        <http://example.org/Various/sf-process/jkl987> ;
      sam:time <http://handle.net/10273/IGSN.SIOabc123/tim2> ] ;
  sam:sampledFeature p1:midAtlanticRidge ;
  sam:samplingFeatureComplex
    [ sam:relatedSamplingFeature <igsn:SIOxyz456> ;
      sam:role p1:parent ] ;
  sam:samplingLocation
    p1:loc123 ;
  sam:samplingMethod <http://ldeo.columbia.edu/sampling/ghostbuster> ;
  sam:samplingTime "2013-06-12T09:25:00.00+11:00"^^xsd:dateTime
  sam:size [ basic:uom unit:kg ;
    basic:value "0.46"^^basic:Number ] ;
  sam:specimenType p1:splitCore .
```
O&M Specimen model

Single model underlies different implementations
Application of O&M

• Direct
  • SOS + OMXML
    – [Metadata for values from] data services
  • Specializations
    – WaterML2, WaterML-WQ, SoilML, SeaDataNet, INSPIRE, WMO ...

• Bridging vocabulary for integration
  • TERENO (GFZ), NEII, ODM2, IGSN/SESAR, OBOE ...

• Checklist ...
  • Slots for implicit → explicit for x-domain use
Modernizing
UML Model

Output in OWL

UML Model

Output in OWL

XSD - OWL from UML
O&M in OWL2

```owl
om:Observation
  a :Class ;
  rdfs:label "Observation"@en ;
  rdfs:subClassOf gf:AnyFeature , h2o:FeatureType ;
  :disjointWith om:Process ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:result ] ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:observedProperty ] ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:featureOfInterest ] ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:phenomenonTime ] ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:procedure ] ;
  rdfs:subClassOf [ a :Restriction ; :cardinality "1"; :onProperty om:resultTime ] .
```

sam:Specimen

Diagram showing the relationships between classes and properties, such as:
- `h2o:FeatureType` with properties like `gml:byValuePropertyType`, `gml:metadataCollection`, and `gml:metadataNoPropertyType`.
- `sam:NotSamplingFeature` connected to `sam:SamplingFeature`.
- `om:Observation` with properties like `om:featureOfInterest`, `om:observedProperty`, and `om:phenomenonTime`.
- `sam:Specimen` with properties like `sam:currentLocation`, `sam:materialClass`, and `sam:samplingLocation`.

Diagram includes a mix of classes and properties indicating the structure of the data model.
SSN ontology
Is Observation a Social Object, or an Event?
Lessons
Lessons

• Ontology for observations, models and forecasts
  • X-domain interoperability
  • Extensible, specialize with vocabularies, link to other ontologies

• Standards
  • A generic model can provide a checklist for design, and a basis for harmonization and incremental design
  • Consensus introduces an overhead

• Change
  • Common conceptual model supports multiple implementations

• Widely used
Additional credits

**O&M**: Fowler & Odell; GeoSciML team; OGC; ISO/TC 211; Rob Atkinson, Rob Woodcock (CSIRO)

**WaterML**: CUAHSI; Gavin Walker, **Pete Taylor**, Laurent Lefort (CSIRO); Paul Sheahan (BOM)

**Soil, WQ**: Bruce Simons, Peter Wilson, Jonathan Yu (CSIRO); Alistair Ritchie (Landcare NZ)

**SeaDataNet**: Jordi Sorribas, Paolo Diviacco

**AQD**: Michael Lutz, Ale Sarretta (JRC); **Kathi Schleidt** (Env. Agency Austria)

**WMO**: Jeremy Tandy (UK Met Office); Aaron Braeckel (UCAR)

**NEII**: Andrew Woolf, Dom Lowe (BOM)

**ODM2**: Jeff Horsburgh

**IGSN**: Kerstin Lehnert (LDEO), Jens Klump (GFZ)

**Vocabulary services**: Stuart Williams (Epimorphics), Roy Lowry, Adam Leadbetter (BODC)
Thank you

CSIRO Land and Water
Simon Cox
Research Scientist

+61 3 9252 6342
simon.cox@csiro.au
www.csiro.au/people/simon.cox