Basic Observations and Sampling Feature Ontology

Editor(s): TBC
Solicited review(s): TBC

Simon J D Cox
CSIRO Land and Water, PO Box 56, Highett, Vic. 3190, Australia
simon.cox@csiro.au

Abstract.

We introduce new OWL ontologies for observations and sampling features, based on the O&M conceptual model from ISO 19156. Previous efforts, through the W3C SSN project, and following the ISO rules for conversion from UML, introduced dependencies on elaborate pre-existing ontologies and frameworks. The new ontologies, known as om-lite and sam-lite, minimize such dependencies, and can therefore be used to harmonize observational data with minimal ontological commitment beyond the conceptual model. Patterns for linking existing ontologies for time and space to stub-classes in the new ontologies are described, thus providing a route for harmonization of more specific observation applications. The PROV-O ontology is re-used to support certain requirements for the description of specimens, and a more general alignment of both observation and sampling feature ontologies with PROV-O is described, as well as mappings to some other observation models and ontologies.

Keywords: observations, sampling, ontology re-use, provenance

1. Introduction

Observations and measurements are used to determine values of properties, though application of some procedure at a particular time and place. The result of an observation is strictly an *estimate* of the true value, conditioned by procedure and circumstances, so description of the latter are important in the assessment of the reliability of the estimate.

A conceptual model for observations and measurements (O&M) is described in ISO 19156:2011 [7,20], which builds on a pattern developed originally by Fowler and O'Dea [11]. The model establishes a domain-neutral vocabulary for an observation and its associated properties. A key design goal was to provide a common terminology for both in-situ observations and remote-sensed observations. This was accomplished by separating concerns, with classes for the feature of interest, the procedure, the observed property, and the act of observation itself. This allows places and times associated with each to be distinct if necessary. O&M also includes an important module for sampling features, covering things like stations, transects, cross-sections, images and specimens. The role of a sampling feature is to assist the characterization of the ultimate feature of interest. They are almost ubiquitous in scientific and environmental observations.

O&M is one of a group of standards developed through Open Geospatial Consortium's Sensor Web Enablement initiative (SWE). O&M provides a user-centric (i.e. user of observation data) viewpoint that complements the provider-centric viewpoint given in SensorML [2]. A GML-based XML implementation of O&M is available for use in the OGC Sensor Observation Service and Web Feature Service [4,18,31].

A number of other projects have developed ontologies for observations. A comprehensive review of the state of the art in 2011 was included in a report from the W3C Semantic Sensor Network incubator group [25], mostly using O&M as a convenient framework for comparing existing observation models and ontologies. The incubator group then developed the Semantic Sensor Network ontology ("SSN") [5], based primarily on the Stimulus-Sensor-Observation pattern (SSO) [22,30], which adds the notion of 'stimulus' into the core O&M model. Meanwhile, Cox [8] developed an ontology for O&M ("OMU") based on automatic conversion of the original UML model, using rules developed in ISO 19150-2 [21].

However, these implementations present barriers to adoption for new applications. In particular

- SSN includes elements for sensors and observations, but omits sampling features, which are a key element required for many practical applications;
- SSN is linked to the Dolce-ultra-lite (DUL) implementation of the DOLCE foundational ontology [12], with SSN concepts directly inheriting from a number of DUL classes and properties. This introduces ontological commitments and a level of complexity which are uncomfortable for some users. For example: Observation is modeled as a sub-class of SocialObject, which is disjoint with Event;
- The UML-OWL conversion rule used for OMU triggers a web of dependencies on additional, sometimes highly detailed ontologies derived from other ISO 19100-series UML models. This introduces a large amount of baggage, of uneven quality, which is unacceptable to some users.

In this paper, we introduce a new OWL implementation of O&M which aims to overcome these limitations with two new ontologies. The new ontologies include both the observation and sampling feature models from O&M, and have fewer dependencies on existing ontologies than OMU and SSN. We expect that these ontologies can either serve as foundations for more domain-specific treatments, or as bridging ontologies for alignment of existing ontologies developed around specific applications or domains.

2. O&M conceptual model

2.1. Observations

The core of the Observation model from O&M [7,20] is shown in UML in Figure 1. Sub-classes of Observation are classified by the result-type, as shown in Figure 2. Note the use of types and classes from other ISO 19100-series models, indicated by the prefixes GF, TM, GM, MD, DQ, LI (from ISO 19109, 19108, 19107 and 19115).

2.2. Sampling features

The core of the Sampling Feature model from O&M [7,20] is shown in Figure 3, and specializations characterized by topological dimension in Figure 4. The O&M model for specimens (i.e. physical samples removed from the natural setting, and used in laboratory observations) is shown in Figure 5.

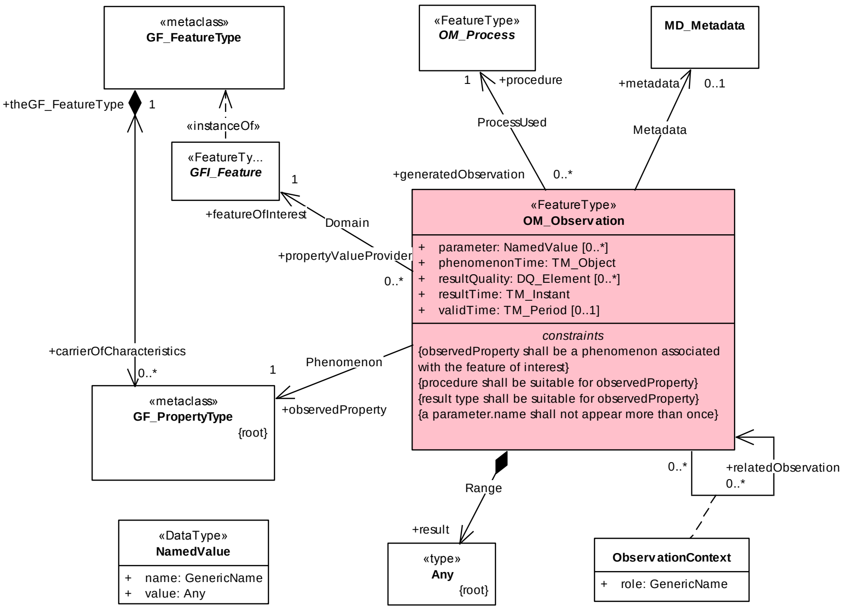

Fig. 1. UML classes and properties in core observation model from ISO 19156:2011

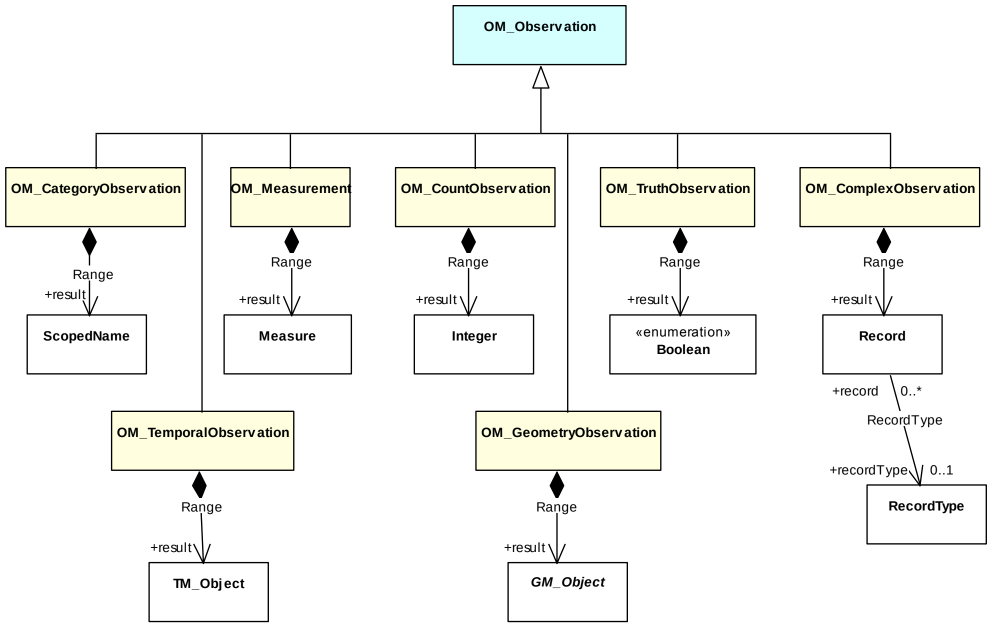

Fig. 2. Observation sub-classes from ISO 19156:2011, characterized by result type.

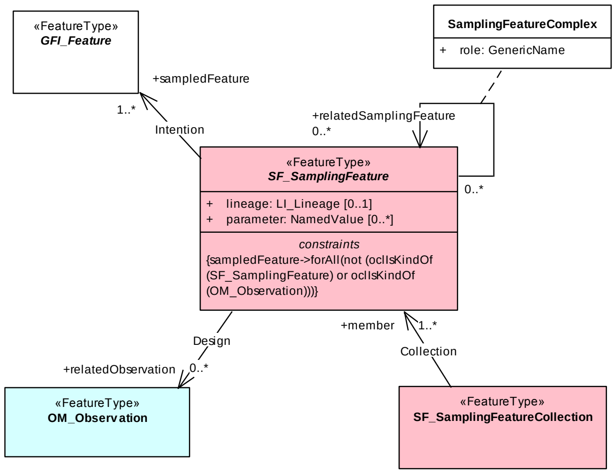

Fig. 3. Classes and properties in core sampling feature model from ISO 19156:2011

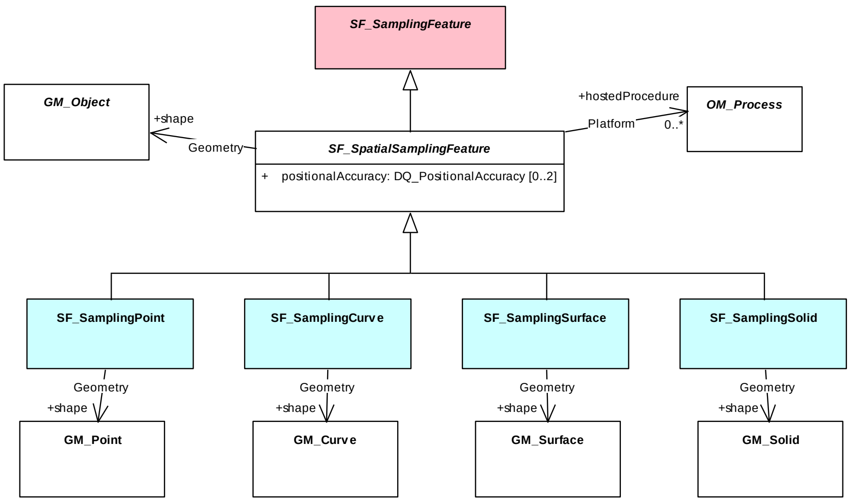

Fig. 4. Sampling feature sub-classes from ISO 19156:2011, characterized by topological dimension.

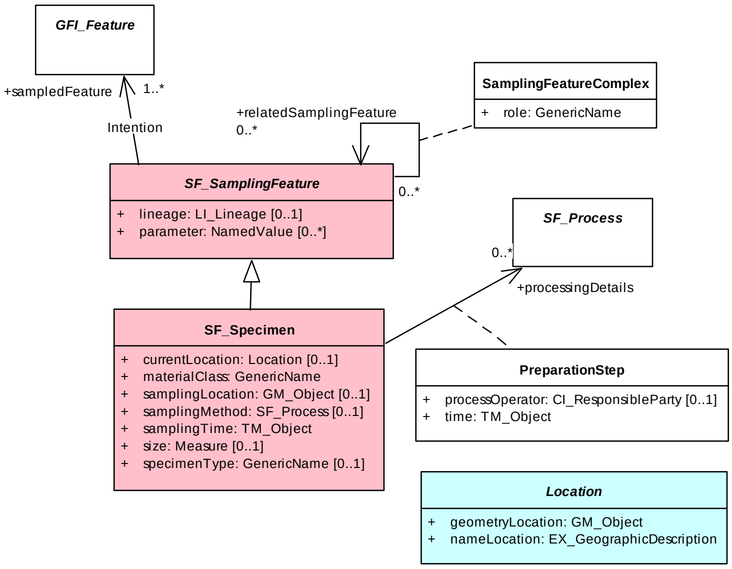

Fig. 5. Specimen model from ISO 19156:2011.

3. OWL implementation

3.1. Observations

The new ontology for observations, known as "om-lite" (namespace prefix "oml:"), covers the key classes from O&M: OM_Observation and its sub-classes, the supporting concept OM_Process, and the association class ObservationContext (Figure 6). Some classes imported from other ISO 19100-series UML models are replaced with local stubs: oml:TemporalObject in place of TM_Object; oml:GeometryObject in place of GM_Object; oml:Measure in place of Measure (for scaled values). Classes are discarded where they provide a subsidiary capability for which well known RDF vocabularies may be used (e.g. the functionality of MD_Metadata is provided by Dublin Core [23,34], DCAT [26], PROV-O [24], etc). The key properties from om-lite are represented as shown in Table 1.

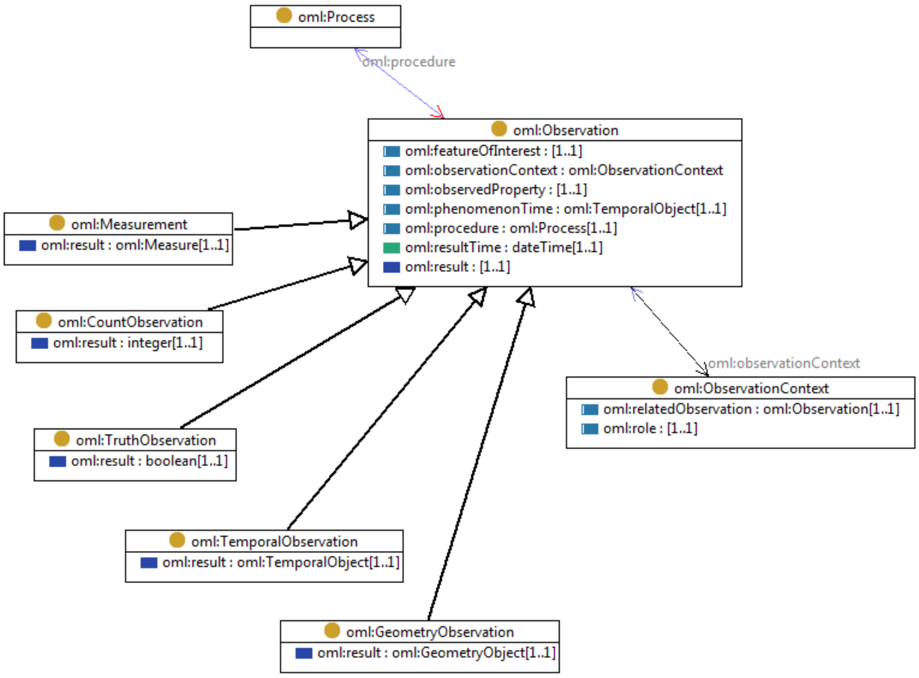

Fig. 6. Basic observation class and specializations. (UML-style view from TopBraid.)

Table 1

Observation properties

Property	**Domain**	**Range**
oml:featureOfInterest	oml:Observation	
oml:observedProperty (functionalProperty)	oml:Observation	
oml:result (functionalProperty)	oml:Observation	
oml:procedure	oml:Observation	oml:Process
oml:phenomenonTime	oml:Observation	oml:TemporalObject
oml:resultTime (functionalProperty)	oml:Observation	xsd:dateTime
oml:observationContext	oml:Observation	oml:ObservationContext
oml:relatedObservation		oml:Observation
oml:role	oml:ObservationContext	

Cardinality restrictions on oml:Observation shown in Figure 6 reflect the expectation that six core properties characterize an observation. The range of oml:resultTime is xsd:dateTime. However, the range of oml:phenomenonTime is oml:TemporalObject, since observations may estimate the value of a property at a wider range of times than is supported by the dateTime datatype (the W3C Time Ontology [13] is also not sufficient, for reasons explained in [9]).

There is no global restriction on the range of oml:featureOfInterest or oml:observedProperty, since a generic model must accommodate observations of any property characterizing any feature or object. Likewise, the range of oml:result is not specified, since property values may have many types and may be characterized in may different ways. A subset of the specialized O&M observation classes is implemented using local restrictions on the type of oml:result.

Note that the [1..1] cardinality restrictions strictly confine users to the OWL 2 DL language profile - the less expressive profiles OWL 2 RL, QL and EL do not permit these restrictions [28].

3.2. Sampling features

The ontology for sampling features, known as "sam-lite" (namespace prefix "saml:") includes all the classes from the O&M Sampling Features model (Figure 7). Following the strategy used in om-lite, stub classes are introduced in the saml: namespace for GeometryObject and its specializations, to serve as the domain of key properties. This allows use of the Sampling Features ontology without incurring the cost of importing a large hierarchy of classes from ontologies implementing ISO/TC 211 models. Table 2 shows the representation in sam-lite of the key properties from O&M sampling features.

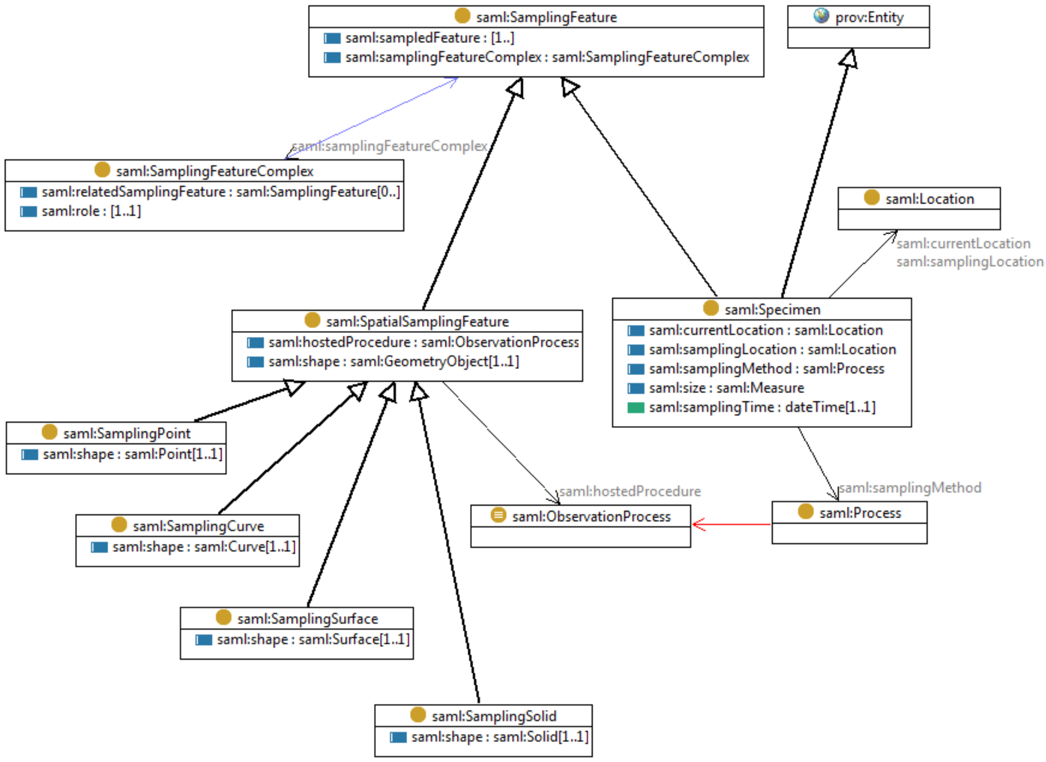

Fig. 7. Sampling features and subclasses

Table 2

Sampling feature properties

Property	Domain	Range
saml:sampledFeature	saml:SamplingFeature	
saml:shape (functionalProperty)	saml:SpatialSamplingFeature	saml:GeometryObject
saml:hostedProcedure	saml:SpatialSamplingFeature	saml:ObservationProcess
saml:samplingFeatureComplex	saml:SamplingFeature	saml:SamplingFeatureComplex
saml:relatedSamplingFeature		saml:SamplingFeature
saml:role	saml:SamplingFeatureComplex	
saml:samplingTime (functionalProperty)	saml:Specimen	xsd:dateTime
saml:samplingMethod	saml:Specimen	saml:Process (disjointWith saml:ObservationProcess)
saml:samplingLocation	saml:Specimen	saml:Location
saml:currentLocation	saml:Specimen	saml:Location
saml:size	saml:Specimen	saml:Measure

A cardinality restriction on saml:SamplingFeature reflects the expectation that a saml:sampledFeature property will be present. No local or global restriction on the range of saml:sampledFeature is provided, since a generic model must accommodate sampling any feature or object.

Sampling of a feature of interest is frequently achieved using a spatially-defined subset. This is represented by the subclass saml:SpatialSamplingFeature, which has a functional property saml:shape, whose range is saml:GeometryObject. Although saml:GeometryObject is equivalent to oml:GeometryObject we do not introduce a dependency on om-lite as there will be applications that only use sampling features. Specific subclasses restrict the type of saml:shape, corresponding to common practice particularly in earth and environmental sciences.

Specimens are physical samples retrieved from their natural environment and used (typically) in laboratory observations. This is represented by a subclass saml:Specimen. Required properties are saml:sampledFeature (from saml:SamplingFeature) and saml:samplingTime. Some additional convenience properties are provided (Table 2).

A critical aspect of specimen description is the record of their preparation and lineage. In the O&M model this was implemented using an association class “PreparationStep”. However, this approach is not fully satisfactory, particularly as the preparation step is not easily linked to an explicit predecessor specimen. In practice there is a very wide range of specimen preparation and provenance paths, so rather than trying to develop a new generic model we have chosen to leverage the W3C PROV ontology [24], which provides patterns for description of relationships between activities, parties and related entities. We make saml:Specimen a subclass of prov:Entity, thus accommodating the requirements of the O&M PreparationStep class as well as relationships with predecessor specimens.

4. Examples

We present a number of examples serialized in Turtle [1]. For these examples the following additional axioms were introduced to allow concrete representations of time and space from existing W3C vocabularies [3,13] to be used as the value of properties whose range is one of the stub classes:

```
<http://www.w3.org/2003/01/geo/wgs84_pos#Point>
    rdfs:subClassOf oml:GeometryObject ;
    rdfs:subClassOf saml:Point ;
.
<http://www.w3.org/2006/time#TemporalEntity>
    rdfs:subClassOf oml:TemporalObject ;
.
```

4.1. Observations

Listing 1 shows a basic measurement of the weight of a piece of fruit, corresponding to the first example in section 5.1 of OMU [8]. The listing is almost the same as the 'explicit' implementation following the ISO 19150-2 rule, but without the dependency on the basic: namespace from ISO 19103 [19] or the tm: namespace from ISO 19108 [17].

Listing 2 shows a remote sensing observation, in which the result is provided as a link to an image dataset, corresponding to C.2.3 from the OGC XML (GML) implementation of O&M [6]. The use of links to resources available elsewhere is natural in the RDF implementation, though some information that is provided in additional xlink attributes alongside the href in the GML implementation is not available locally in the RDF.

4.2. Sampling features and specimens

Listing 3 shows the description of a river sampling station including links to two observations made there, corresponding to C.3.1 from the OGC XML (GML) implementation of O&M [6]. The station is a member of a collection of sampling features. This membership is captured using the saml:samplingFeatureComplex property. We have also shown how this may be captured using prov:wasMemberOf from the PROV ontology [24].

Listing 4 shows a description of a specimen of rock, corresponding to the example in section 5.2 of [8]. The description includes links to the parent specimen from which it was generated. The property prov:wasGeneratedBy links to the activity that generated the current specimen, which in turn links to the previous specimen, the process operator, and to the processing method. This example illustrates how the standard PROV ontology supports typical specimen preparation metadata directly. More detailed models could be implemented by further specialization of the PROV-O properties.

Listing 1 – A simple measurement example

```
my:obsTest1  a                oml:Measurement ;
        rdfs:comment          "Observation test instance: fruit mass"^^xsd:string ;
        rdfs:label            "Observation test 1"^^xsd:string ;
        oml:featureOfInterest  <http://wfs.example.org?request=getFeature&amp;featureid=fruit37f> ;
        oml:observedProperty   <http://sweet.jpl.nasa.gov/2.0/phys.owl#Mass> ;
        oml:phenomenonTime     [ a                        w3time:Instant ;
                                 w3time:inXSDDateTime  "2005-01-11T16:22:25.00"^^xsd:dateTime
                              ] ;
        oml:procedure          my:Sscales1 ;
        oml:result             [ a          oml:Measure ;
                                 rdf:value  "0.28"^^oml:Number ;
                                 oml:uom     <http://www.opengis.net/def/uom/UCUM/0/kg>
                              ] ;
        oml:resultTime         "2005-01-11T16:22:25.00"^^xsd:dateTime .
```

Listing 2 – An observation whose result is provided out-of-band

```
my:OPTest1  a                 oml:Observation ;
        rdfs:comment          "Observation instance with remote result"^^xsd:string ;
        rdfs:label            "Observation Pointer 1"^^xsd:string ;
        oml:featureOfInterest  <http://my.example.org/wfs%26request=getFeature%26;featureid=789002> ;
        oml:observedProperty   <http://vocab.nerc.ac.uk/collection/I01/current/0.1.1/> ;
        oml:phenomenonTime     [ a                      w3time:ProperInterval ;
                                 w3time:hasBeginning  [ a                     w3time:Instant ;
                                                        w3time:inXSDDateTime  "2005-01-11T17:22:25.00"^^xsd:dateTime
                                                      ] ;
                                 w3time:hasEnd        [ a                     w3time:Instant ;
                                                        w3time:inXSDDateTime  "2005-01-11T18:22:25.00"^^xsd:dateTime
                                                      ]
                              ] ;
        oml:result             <http://my.example.org/results%3f798002%26property=RH> ;
        oml:resultTime         "2005-01-11T18:22:25.00"^^xsd:dateTime .
```

Listing 3 – A sampling station with links to some related observations

```
<http://my.hydrology.example.org/catchments/Potamos#st2>  a  saml:SamplingPoint ;
        rdfs:comment                "Hydrology sampling station"^^xsd:string ;
        oml:relatedObservation       <http://my.hydrology.example.org/chemistry/2007/rtg78n> ,
                                    <http://my.hydrology.example.org/chemistry/2007/rtg108q> ;
        saml:sampledFeature          <http://my.hydrology.example.org/catchments/Potamos> ;
        saml:samplingFeatureComplex  [ a          saml:SamplingFeatureComplex ;
                                      saml:relatedSamplingFeature
                                                 <http://my.example.org/wfs?request=getFeature;featureid=coll32> ;
                                      saml:role   <http://www.example.org/complex/member>
                                    ] ;
        prov:wasMemberOf            <http://my.example.org/wfs?request=getFeature;featureid=coll32> .
```

Listing 4 – A specimen with provenance and preparation information

```
<http://handle.net/10273/IGSN.SIOabc123>  a  saml:Specimen , my:splitCore ,
                                         <http://www.opengis.net/def/material/OGC-OM/2.0/rock> ;
        rdfs:label                  "SIO specimen abc123"^^xsd:string ;
        saml:sampledFeature          my:midAtlanticRidge ;
        saml:samplingMethod          <http://ldeo.columbia.edu/sampling/ghostbuster> ;
        saml:samplingTime            "2013-06-12T09:25:00.00+11:00"^^xsd:dateTime ;
        saml:samplingLocation        [ a           w3geo:Point ;
                                      w3geo:alt   -1272.0 ;
                                      w3geo:lat   24.97 ;
                                      w3geo:long  -45.87
                                    ] ;
        saml:currentLocation         <http://example.org/various/Warehouse3/shelf9/box67> ;
        saml:size                    [ a          saml:Measure ;
                                      rdf:value  "0.46"^^saml:Number ;
                                      saml:uom    <http://qudt.org/vocab/unit#Kilogram>
                                    ] ;
        saml:samplingFeatureComplex  [ a                           saml:SamplingFeatureComplex ;
                                      saml:relatedSamplingFeature  <http://handle.net/10273/IGSN.SIOxyz456> ;
                                      saml:role                    my:parent
                                    ] ;
        prov:wasDerivedFrom         <http://handle.net/10273/IGSN.SIOxyz456> ;
        prov:wasGeneratedBy         [ a                     prov:Activity ;
                                      prov:endedAtTime        "2013-08-02T08:15:00.00+11:00"^^xsd:dateTime ;
                                      prov:used               <http://handle.net/10273/IGSN.SIOxyz456> ;
                                      prov:wasAssociatedWith  my:JohnDoe ;
                                      prov:wasInformedBy      <http://example.org/various/sf-process/jkl987>
                                    ] .
```

5. Discussion

5.1. Dependencies

In Figure 8 we show the Observation class and its dependency classes as expressed in om-lite, SSN and OMU. From this view it is clear that om-lite does provides a light-weight framework, in comparison with both SSN and OMU which any application to make a significant commitment to an existing framework.

5.2. Alignment

5.2.1. Time and space

Stub classes are introduced in om-lite and sam-lite to implement time and space concepts used in the O&M observation and sampling feature models. These classes are required to support the definition of classes which have constraints involving these types, but the new classes do not commit the user to any specific existing model or ontology for geometric or temporal objects. When the ontology is used for data individuals, additional axioms must be introduced (or will be inferred) that link the classes from the stub classes to some concrete representation of time and space. As anticipated in section 0, this may be done by locally asserting that the type used is a sub-class of the stub class, which is therefore understood to be the superclass of all possible representations of that concept.

Figure 9 shows a set of possible sub-class relationships linking from the W3C Basic Geo [3], W3C Time [13], GeoSPARQL [29] and ISO 19150-2 ontologies [8,21] to the stub classes introduced in om-lite and sam-lite. Since the stub classes have no properties or constraints (or further superclasses) these subclassing axioms are "conservative" and thus non-harmful in the sense described by Hogan et al. [14,15]. Nevertheless, they should only be introduced locally, in the context of individual data instances, because adding superclasses to legacy classes generally is inadvisable, with ontology 'hijacking' risking both performance and reasoning behaviour [14,15].

5.2.2. SSN ontology

The SSN ontology [5] can be aligned with the om-lite ontology using a similar approach: classes from the application may be asserted to be sub-classed from the equivalent classes in om-lite. Listing 5 shows possible sub-class and sub-property relationships linking the ontologies. Since om-lite is not aligned to any particular foundational ontology, SSN's dependency on DOLCE should not introduce conflicts.

Note that sam-lite provides an ontology for sampling features that is missing from SSN.

5.2.3. Domain ontologies

Ontologies for observation applications may use one of two approaches to align with om-lite and sam-lite.

1. A new ontology may be explicitly based on om-lite and sam-lite. Classes and properties from the application ontology can be used as-is where suitable. Other classes may be specialized from the om-lite and sam-lite ontologies, adopting the proposed axioms and inheriting the existing constraints, or may be added with new relationships.
2. An existing ontology can be aligned with om-lite and sam-lite by asserting class-class and property-property relationships as in the examples above. For example: Listing 6 shows relationships to align the OBOE ontology [27] with om-lite; Listing 7 shows relationships to align ODM2 [16] with om-lite, sam-lite and PROV-O (ODM2 resource names inferred from UML).

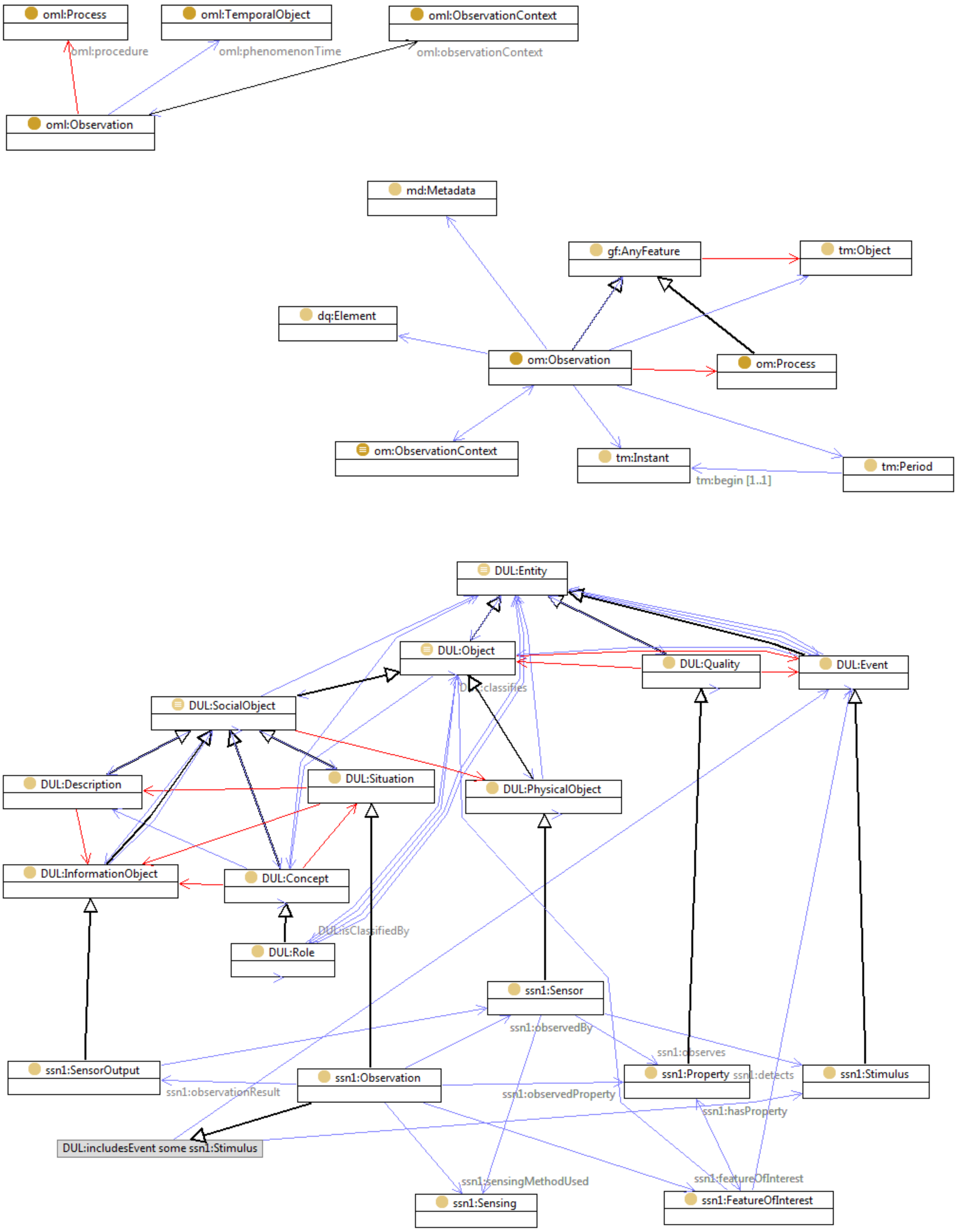

Figure 8 – Comparison of the Observation class and its dependencies in the om-lite (top), OMU (middle) and SSN (bottom) ontologies

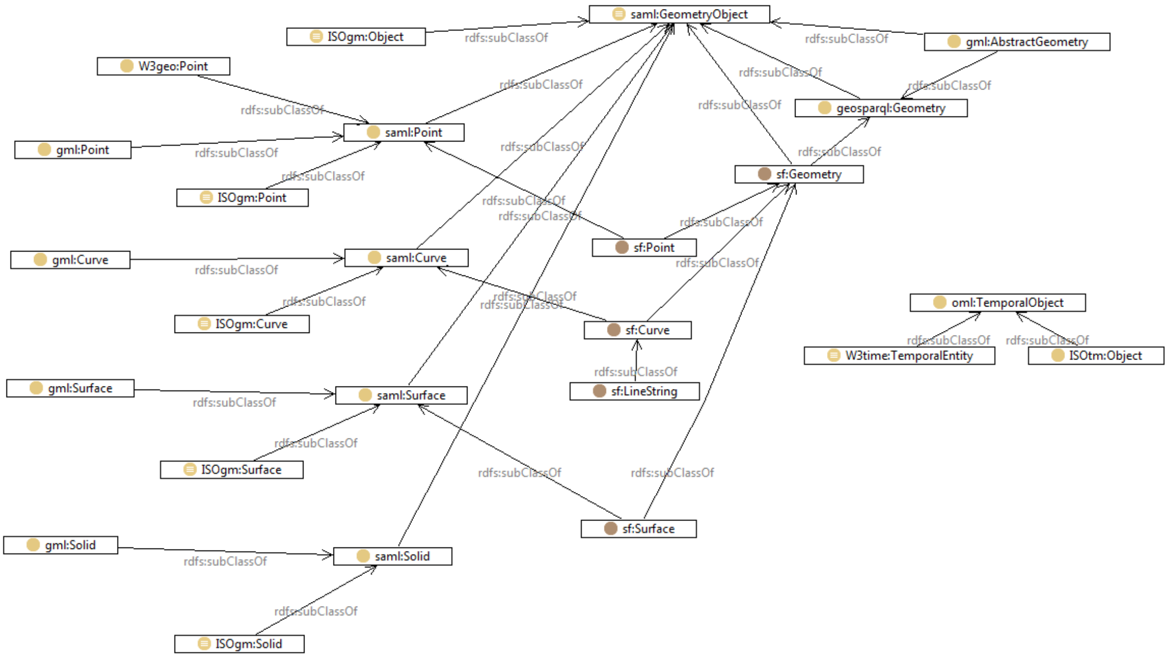

Figure 9 – Alignment of existing ontologies with stub classes for space and time from om-lite and sam-lite.

Listing 5 – Alignment of classes and properties from SSN [5] with om-lite

```
ssn:Observation             rdfs:subClassOf             oml:Observation .
ssn:Sensing                 rdfs:subClassOf             oml:Process .
ssn:Sensor                  rdfs:subClassOf             oml:Process .
ssn:featureOfInterest       rdfs:subPropertyOf          oml:featureOfInterest .
ssn:observationResult       rdfs:subPropertyOf          oml:result .
ssn:observationResultTime   rdfs:subPropertyOf          oml:resultTime .
ssn:observedBy              rdfs:subPropertyOf          oml:procedure .
ssn:observedProperty        rdfs:subPropertyOf          oml:observedProperty .
ssn:sensingMethodUsed       rdfs:subPropertyOf          oml:procedure .
```

Listing 6 – Alignment of classes and properties from OBOE [27] with om-lite

```
oboe-core:Measurement       rdfs:subClassOf             oml:Observation .
oboe-core:Protocol          rdfs:subClassOf             oml:Process .
oboe-core:hasContext        rdfs:subPropertyOf          oml:observationContext .
oboe-core:hasValue          rdfs:subPropertyOf          oml:result .
oboe-core:ofCharacteristic  rdfs:subPropertyOf          oml:observedProperty .
oboe-core:ofEntity          rdfs:subPropertyOf          oml:featureOfInterest .
oboe-core:usesMethod        rdfs:subPropertyOf          oml:procedure .
oboe-core:usesProtocol      rdfs:subPropertyOf          oml:procedure .
```

Listing 7 – Alignment of classes and properties from ODM2 [16] with om-lite, sam-lite and PROV-O.

```
odm2:Action                 rdfs:subClassOf             prov:Activity .
odm2:ObservationAction      rdfs:subClassOf             oml:Observation .
odm2:Organization           rdfs:subClassOf             prov:Agent .
odm2:Person                 rdfs:subClassOf             prov:Agent .
odm2:SamplingFeature        rdfs:subClassOf             saml:SamplingFeature .
odm2:featureOfInterest      rdfs:subPropertyOf          oml:featureOfInterest .
odm2:result                 rdfs:subPropertyOf          oml:result .
odm2:variable               rdfs:subPropertyOf          oml:observedProperty .
```

5.3. Use of W3C PROV-O ontology

5.3.1. Sampling features ontology

PROV-O [24] is the only legacy ontology used directly in the new ontologies, apart from the basic RDF, RDFS and OWL infrastructure. The primary axiom linking sam-lite to PROV-O is

```
saml:Specimen rdfs:subClassOf prov:Entity .
```

The motivation for the introduction of PROV-O in sam-lite was to support flexible description of specimen preparation chains, the details of which vary widely in different disciplines and communities. PROV-O provides a large set of generic properties to support relationships between Entities, Activities and Agents, and appears to be well scoped to this task.

However, as well as replacing the PreparationStep property in the context of specimens, in Listings 3 and 4 above we show that PROV-O relationships could be used in place of at least some applications of the samplingFeatureComplex property from the Sampling Features model. The statement involving prov:wasMemberOf [1] in Listing 3 entails that saml:SamplingPoint is either equivalent to or subclassed from prov:Entity. Thus, a more general subclassing axiom might be introduced:

```
saml:SamplingFeature rdfs:subClassOf prov:Entity .
```

The diversity of potential relationships between sampling features within a complex was managed in the original O&M model through a "role" property on the SamplingFeatureComplex class, which is implemented directly in sam-lite (Table 2). In order to take advantage of the general subclassing relationship proposed above, the functionality of "role" would be implemented by sub-properties of PROV-O properties whose domain and range allow for prov:Entity, in particular prov:wasInfluencedBy and prov:wasDerivedFrom. However, this cannot capture the semantics for some common relationships between sampling features in a complex (such as the geometric relationships that connect stations to a transect, pixels to an image, or specimens to a borehole) because they are "sibling" relationships rather than derivation relationships. To retain the general functionality we still need a property for non-derivation relationships between sampling features.

[1] "wasMemberOf" is not a formal PROV-O property, but is a reserved name recommended for the inverse of prov:hadMember.

Another potential subclassing opportunity in sam-lite is

```
saml:Process rdfs:subClassOf prov:Entity .
```

prov:Entity is preferred over sub-classing prov:Activity, since saml:Process and prov:Entity are continuants, while a prov:Activity is generally understood to be an occurrent.

5.3.2. Observation ontology

In O&M, the description of an Observation is understood primarily as providing provenance for its result. The following axioms interpret om-lite in terms of PROV-O:

```
oml:Observation rdfs:subClassOf    prov:Activity .
oml:resultTime  rdfs:subPropertyOf prov:endedAtTime .
oml:result      rdfs:subPropertyOf prov:generated .
oml:result      rdfs:range         prov:Entity .
oml:procedure   rdfs:subPropertyOf prov:used .
oml:Process     rdfs:subClassOf    prov:Entity .
```

However, we have not included this alignment with PROV-O as an explicit part of om-lite at this time, since a primary design goal was to minimize dependencies.

5.3.3. Information resources or real-world things8

The PROV-O specification [24] is clear that the intention is for PROV-O be applicable to things in the real world. However, the examples in the standard use prov:Entity almost exclusively for information resources (reports, documents, datasets, graphs). One minor example of a biological specimen (drosophila-a) is mentioned in PROV-O, but has a very short provenance chain.

In the alignments proposed here, prov:Entity is the superclass for saml:Specimen, saml:SamplingFeature, saml:Process and oml:Process, which are either physical or virtual objects in the world, and process types, and not just documents. We thus demonstrate the applicability of PROV-O to real-world things.

5.4. Sampling features

Finally, it is notable that few of the other observation models and ontologies in use make the role of sampling features in the observation process explicit. This is somewhat surprising, as sampling is ubiquitous in practical observations scenarios. Sampling involves subsetting the ultimate feature of interest in some way, and it is helpful to identify and describe both sampling features and the ultimate feature of interest separately and explicitly.

Spatially defined sampling is common in multiple domains in earth and environmental sciences (with feature names like station, transect, cross-section, swath etc), and multiple features are typically linked within a sampling strategy (specimens along a borehole; stations on a transect; flight-lines within an aerial survey; pixels within an image). The O&M spatial sampling features model was particularly influenced by Climate Science Modelling Language from the 'fluid-earth' community (oceans and atmospheres) [33], and the specimen model was influenced by a wider variety of use-cases, particularly geochemistry [10] and work in the biodiversity community that led on to the development of the Biological Collections Ontology (BCO) [32].

The sampling features model in O&M provides a kernel for direct use or domain-based extension, and its implementation in sam-lite is thus a very important component of the observation ontologies

6. Summary

We have described basic OWL ontologies for observations and for sampling features, which implement the concepts from the ISO O&M model. Unlike previous attempts, these new ontologies have no dependencies on elaborate ontology networks or foundational ontologies, and thus do not require the user to commit to any other existing framework.

The single exception is the model for specimens, which re-uses elements from PROV-O to overcome some known limitations of the O&M model. We describe other potential alignments with PROV-O, which is particularly appropriate for the observation model, whose goal is to provide structured provenance information for estimates of property values.

The new ontologies may be used as-is[2], but are likely to be of more value in providing a basis for more specialized and application specific observation ontologies, or as a bridging ontology to assist in linking between existing models or ontologies.

Acknowledgements

Thanks to David Ratcliffe for helpful pointers on OWL details as usual, and Jonathan Yu for assistance in highlighting the key messages. This work is a contribution towards the harmonization of W3C and Open Geospatial Consortium standards for geospatial data, and was supported by the CSIRO Land and Water Flagship, the Water Information Research and Development Alliance, and the eReefs project.

[2] The om-lite ontology is published at http://def.seegrid.csiro.au/ontology/om/om-lite . The sam-lite ontology is published at http://def.seegrid.csiro.au/ontology/om/sam-lite .

References

[1] D. Beckett, T. Berners-Lee, E. Prud'hommeaux, G. Carothers, RDF 1.1 Turtle, World Wide Web Consortium, 2014, Available at: http://www.w3.org/TR/2014/REC-turtle-20140225/.

[2] M. Botts, Sensor Model Language (SensorML) | OGC, Open Geospatial Consortium, Wayland, Mass., 2013, Available at: http://www.opengeospatial.org/standards/sensorml.

[3] D. Brickley, W.S.W.I. Group, Basic Geo (WGS84 lat/long) Vocabulary, 2004, Available at: http://www.w3.org/2003/01/geo/.

[4] A. Bröring, C. Stasch, J. Echterhoff, Sensor Observation Service | OGC, Open Geospatial Consortium, Wayland, Mass., 2012, Available at: http://www.opengeospatial.org/standards/sos.

[5] M. Compton, P. Barnaghi, L. Bermudez, R. García-Castro, O. Corcho, S.J.D. Cox, et al., The SSN ontology of the W3C semantic sensor network incubator group, Web Semant. Sci. Serv. Agents World Wide Web. 17 (2012) 25–32 DOI:10.1016/j.websem.2012.05.003.

[6] S.J.D. Cox, Observations and Measurements - XML Implementation, 2.0 ed., Open Geospatial Consortium, Wayland, Mass., 2011, Available at: http://portal.opengeospatial.org/files/41510.

[7] S.J.D. Cox, Geographic Information - Observations and Measurements (OGC Abstract Specification Topic 20), 2.0 ed., Open Geospatial Consortium, Wayland, Mass., 2013, Available at: http://portal.opengeospatial.org/files/41579.

[8] S.J.D. Cox, An explicit OWL representation of ISO/OGC Observations and Measurements, in: O. Corcho, C. Henson, P. Barnaghi (Eds.), Proc. 6th Int. Work. Semant. Sens. Networks Co-Located with 12th Int. Semant. Web Conf. (ISWC 2013), Sun SITE Central Europe, Sydney, Australia, October 22nd, 2013., 2013: pp. 1–18.

[9] S.J.D. Cox, Time Ontology Extended for Non-Gregorian Calendar Applications, Semant. Web J. submitted (2014).

[10] S.J.D. Cox, A. Dent, S. Girvan, R. Atkinson, I. Whitehouse, C. Legg, Using the Assay Data Exchange standard with WFS to build a complete minerals exploration data-transfer chain, in: Proc. Int. Assoc. Math. Geol. Gen. Assem., 2006.

[11] M. Fowler, Analysis Patterns: Reusable Object Models, Addison-Wesley, 1997, Available at: http://martinfowler.com/books/ap.html.

[12] A. Gangemi, Ontology:DOLCE+DnS Ultralite - Odp. [Online]. Available at: http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+DnS_Ultralite. [Accessed: 13-Feb-2014].

[13] J.R. Hobbs, F. Pan, Time Ontology in OWL, World Wide Web Consortium, 2006, Available at: http://www.w3.org/TR/owl-time/.

[14] A. Hogan, A. Harth, A. Passant, S. Decker, A. Polleres, Weaving the Pedantic Web, in: Linked Data Web Work., CEUR Workshop Proceedings, Raleigh, North Carolina, USA, April 27, 2010, 2010.
[15] A. Hogan, A. Harth, A. Polleres, SAOR: Authoritative Reasoning for the Web, Int. J. Semant. Web Inf. Syst. 5 (2009) DOI:10.1007/978-3-540-89704-0_6.
[16] J.S. Horsburgh, A.K. Aufdenkampe, K.A. Lehnert, E. Mayorga, D.G. Tarboton, I. Zaslavsky, et al., Extending the Interoperability of Sensor and Sample Based Earth Observations Using a Community Information Model, in: AGU Fall Meet., San Francisco, 2013.
[17] ISO/TC-211, ISO 19108:2002 - Geographic information -- Temporal schema, International Organization for Standardization, Geneva, 2002, Available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26013.
[18] ISO/TC-211, ISO 19115:2003 - Geographic information -- Metadata, International Organization for Standardization, Geneva, 2003, Available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020.
[19] ISO/TC-211, ISO/TS 19103:2005 - Geographic information -- Conceptual schema language, International Organization for Standardization, Geneva, 2005, Available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=37800.
[20] ISO/TC-211, ISO 19156:2011 - Geographic information -- Observations and measurements, International Organization for Standardization, Geneva, 2011, Available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=32574.
[21] ISO/TC-211, ISO/DIS 19150-2:2014 - Geographic information -- Ontology -- Part 2: Rules for developing ontologies in the Web Ontology Language (OWL), International Organization for Standardization, Geneva, 2014, Available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=57466.
[22] K. Janowicz, M. Compton, The Stimulus-Sensor-Observation Ontology Design Pattern and its Integration into the Semantic Sensor Network Ontology, in: 3rd Int. Work. Semant. Sens. Networks, 2010.
[23] J. Kunze, T. Baker, The Dublin Core Metadata Element Set, IETF RFC. 5013 (2007).
[24] T. Lebo, S. Sahoo, D. McGuinness, PROV-O: The PROV Ontology. [Online]. Available at: http://www.w3.org/TR/prov-o/. [Accessed: 13-Feb-2014].
[25] L. Lefort, C. Henson, K. Taylor, P. Barnaghi, M. Compton, O. Corcho, et al., Incubator Report - Semantic Sensor Network Incubator Group, W3C, 2011, Available at: http://www.w3.org/2005/Incubator/ssn/wiki/Incubator_Report.
[26] F. Maali, J. Erickson, Data Catalog Vocabulary (DCAT), Cambridge, Mass. USA, 2014, Available at: http://www.w3.org/TR/vocab-dcat/.
[27] J. Madin, S. Bowers, M. Schildhauer, S. Krivov, D. Pennington, F. Villa, An ontology for describing and synthesizing ecological observation data, Ecol. Inform. 2 (2007) 279–296 DOI:10.1016/j.ecoinf.2007.05.004.
[28] B. Motik, B.C. Grau, I. Horrocks, Z. Wu, A. Fokoue, C. Lutz, OWL 2 Web Ontology Language Profiles (Second Edition), Cambridge, Mass. USA, 2012, Available at: http://www.w3.org/TR/owl2-profiles/.
[29] M. Perry, J. Herring, OGC GeoSPARQL – a geographic query language for RDF data., Open Geospatial Consortium, 2012, Available at: https://portal.opengeospatial.org/files/47664 .
[30] C. Stasch, K. Janowicz, A. Bröring, I. Reis, W. Kuhn, A stimulus-centric algebraic approach to sensors and observations, in: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2009: pp. 169–179.
[31] P. Vretanos, Web Feature Service | OGC Document 09-025r1, 2.0 ed., Open Geospatial Consortium Inc., Wayland, Mass., 2010, Available at: http://www.opengeospatial.org/standards/wfs.
[32] R.L. Walls, J. Deck, R. Guralnick, S. Baskauf, R. Beaman, S. Blum, et al., Semantics in support of biodiversity knowledge discovery: an introduction to the biological collections ontology and related ontologies., PLoS One. 9 (2014) e89606 DOI:10.1371/journal.pone.0089606.
[33] A. Woolf, B. Lawrence, R. Lowry, K. Kleese van Dam, R. Cramer, M. Gutierrez, et al., Data integration with the Climate Science Modelling Language, Adv. Geosci. 8 (2006) 83–90 DOI:10.5194/adgeo-8-83-2006.
[34] Dublin Core Metadata Element Set, Version 1.1. [Online]. Available at: http://dublincore.org/documents/dces/. [Accessed: 30-Mar-2014].