Open
Geospatial
Consortium

OGC TESTBED-17: SIF SEMANTIC MODEL ENGINEERING REPORT

ENGINEERING REPORT

PUBLISHED

Note

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

I ABSTRACT

This Engineering Report (ER) presents an analysis of the semantic model of the Sensor Integration Framework (SIF). After reviewing the current SIF Semantic Model, existing related ontologies are reviewed. The ER discusses the results and includes all lessons learned from the experiments completed by the Sensor Integration thread of the OGC Testbed-17 initiative. The ER presents a series of recommendations based on the lessons learned.

II EXECUTIVE SUMMARY

A significant barrier to sensor integration is often the variety of standards, formats, and protocols employed in sensor systems. To build impactful sensor systems, it is necessary for such systems to embrace this diversity. There is, therefore, a need for a framework of standards that facilitates sensor integration independently of technological restrictions.

This Engineering Report (ER) presents an analysis of the Semantic Model of the Sensor Integration Framework (SIF). The purpose of the SIF is to provide the guidance required for sensor data producers and consumers to implement a sensor information enterprise that meets operational requirements.

The following future work items have been identified:

- Multi-level Sensor Ontology Architecture: Define a best practice architecture to split the scope and responsibility of different levels of sensor and observable semantics into several ontologies maintained by different organizations, to provide more flexibility when concepts need to be added.

- Cross-Domain Sensor Ontology: Develop an OGC cross-domain ontology that includes concepts that are applicable to many sensor types. It could be hosted on the OGC Definitions Server.

- More Formal Mapping Definitions: One may want to formally define the mathematical relationships between some of the concepts. More advanced mapping ontologies could be developed to that end.

III KEYWORDS

The following are keywords to be used by search engines and document catalogues.

ogcdoc, OGC document, SIF, semantic, sensor, ontology

IV PREFACE

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

V SECURITY CONSIDERATIONS

No security considerations have been made for this document.

VI SUBMITTING ORGANIZATIONS

The following organizations submitted this Document to the Open Geospatial Consortium (OGC):

- University of Calgary
- Botts Innovative Research

VII SUBMITTERS

All questions regarding this document should be directed to the editor or the contributors:

NAME	ORGANIZATION	ROLE
Mahnoush Alsadat Mohammadi Jahromi	University of Calgary	Co-editor and Contributor
Alex Robin	Botts Innovative Research	Co-editor and Contributor
Charles Heazel	Heazel Technologies	Contributor

1

SCOPE

1 SCOPE

This Engineering Report (ER) presents an analysis of the semantic model of the Sensor Integration Framework (SIF). This ER consists of the following chapters:

- Chapter 4 explains why semantics are required.

- Chapter 5 discusses the current SIF semantic model.

- Chapter 6 discusses the implementation of the SIF semantic model.

- Chapter 7 discuss the issues and recommendations identified through this testbed.

- Chapter 8 presents recommendations for future work.

2

TERMS, DEFINITIONS AND ABBREVIATED TERMS

─────

2 TERMS, DEFINITIONS AND ABBREVIATED TERMS

This document uses the terms defined in OGC Policy Directive 49, which is based on the ISO/IEC Directives, Part 2, Rules for the structure and drafting of International Standards. In particular, the word "shall" (not "must") is the verb form used to indicate a requirement to be strictly followed to conform to this document and OGC documents do not use the equivalent phrases in the ISO/IEC Directives, Part 2.

This document also uses terms defined in the OGC Standard for Modular specifications (OGC 08-131r3), also known as the 'ModSpec'. The definitions of terms such as standard, specification, requirement, and conformance test are provided in the ModSpec.

For the purposes of this document, the following additional terms and definitions apply.

2.1. Terms and definitions

2.1.1. **Reference View**

Within the context of the SIF, this is the abstract architecture framework.

2.1.2. **Technical View**

Within the context of the SIF, this provides instruction on the implementation of SIF in a specific environment.

2.2. Abbreviated terms

BIIF   Basic Image Interchange Format

DDIL   Denied, Degraded, Intermittent, or Limited Bandwidth

DoD	Department of Defense
DoDAF	DoD Architecture Framework
FoI	Feature of Interest
GWG	Geospatial-Intelligence Standards Working Group
GWS	Geospatial Web Services
ISA	Integrated Sensor Architecture
ITU-T	Telecommunication Standardization Sector of the International Telecommunications Union
KLV	Key-Length-Value
MIDSI	Motion-Imagery-Derived Still Imagery
MISB	Motion Imagery Standards Board
NITF	National Imagery Transmission Format
NSG	National System for Geospatial-Intelligence
NTB	NITF Technical Board
ODNI	Office of the Director of National Intelligence
OGC	Open Geospatial Consortium
ORM	OGC Reference Model
OSH	OpenSensorHub
OWL	Web Ontology Language
RM-ODP	Reference Model for Open Distributed Processing
RV	Reference View
SIF	Sensor Integration Framework
SMPTE	Society of Motion Picture and Television Engineers
SSN	Semantic Sensor Network
SWE	OGC Sensor Web Enablement
TV	Technical View
UAS	Unmanned Air System
UML	Unified Modeling Language

W3C   World Wide Web Consortium

3

INTRODUCTION

3 INTRODUCTION

The Sensor Integration Framework (SIF) provides a framework, a mediation service or services, to authorized users to integrate different sensors and access their data regardless of any concerns related to their technical constraints, such as sensor size, weight, power, communications capabilities, etc. The SIF proposes a set of Technical Views (TVs) to cover all deployment environments and their associated constraints, and a Reference View (RV). The semantics of the SIF are based on the Semantic Sensor Network (SSN) Ontology which facilitates the information flow between different TVs. The SIF also applies aspects of the OGC Sensor Web Enablement (SWE) suite of standards. SWE is an infrastructure that enables users to easily share their sensor resources in a well-defined way (OGC 13-032).

In this ER, the SIF Semantic Model is reviewed, its potential to grow into a comprehensive semantic model for sensor data is analyzed, and recommendations on if and how it should evolve into an OGC resource are provided.

4

WHY SEMANTICS ARE REQUIRED

4 WHY SEMANTICS ARE REQUIRED

The SIF seeks to enable the seamless flow of sensor information to wherever it is needed. A simple solution would be to mandate a common set of data standards and protocols for all sensor systems. A simple solution that will not work.

The primary goal of a sensor system is to measure a phenomenon wherever it occurs. This often places severe constraints on sensors and sensor systems. These constraints, in turn, impose technology and operational constraints within which the sensor systems have to perform. No single set of standards can address all possible constraints. The SIF addresses this issue by standardizing, not sensor systems, but how information is exchanged between sensor systems. These interfaces are addressed at two levels:

- At the technology level, the SIF addresses protocol transformation and data format transformation. These technologies are relatively mature.

- The second level deals with the information interface. It's one thing to pass a floating-point number between two systems. It's another to convey what that number means. The semantic model addresses how meaning is passed between sensor systems.

The SIF leverages the OGC's Sensor Web Enablement (SWE) Common Standard (OGC 08-094r1), and the SSN Ontology that has been developed jointly by W3C and OGC (OGC 16-079). SWE Common provides a standard set of data types, sufficient for most data that needs to be exchanged between sensor systems. This would enable data transformation by mapping a local data model into the corresponding SWE Common representation. But the semantic definition of the data element is intentionally omitted.

SWE Common entities have a `definition` attribute. This attribute allows us to associate an ontology with an SWE Common dataset. So the definition attribute points to a concept in an ontology that provides the meaning of that data element. If the developer of a sensor system captures their information model in an ontology, then we have the information fully defined both semantically (the ontology) and syntactically (SWE Common).

4.1. Challenge 1

One challenge is transforming from one ontology to another which will be covered in Testbed-17 and 18. Comparing two different ontologies would become easier if their corresponding semantic model contains different abstraction levels of knowledge. In a URI-based semantic model, for example, where at the top there are the most abstract, and as it goes to the bottom, more specific semantics are provided. At the time of comparing two different ontologies, it would be limited to comparing two URIs at a very high level.

The semantic model is essentially a platform-independent model at the top down, and it is an extension of the SSN Ontology and SWE Common data model, where other ontologies, each specific to a particular sensor system, can map into.

MathML (Mathematical Markup Language) was once proposed in OGC, which allows expressing formulas and equations, but since there were not sufficient tools at the time to load MathML, the idea was dropped. In SIF, only concepts that have multiple ways of mathematically achieving the same behavior are considered, for example, the concept of rotation which is discussed further in this document.

The ideal situation and the goal here is the stage where sensor systems such as MASBUS and OpenSensorHub, for example, can semantically translate on-the-fly. Therefore, for a new sensor, as an example, OpenSensorHub will convert all the sensor properties and parameters into SIF, and if it is equivalent, then it will provide a single comment model to the user.

4.2. Challenge 2

How does the SIF semantic model support new points of view to observations? How would the SIF help a user to compare two different Features of Interest (FoI) to each other?

For example, in a scenario where the number of specific groups of people in the city matters, we can have multiple interpretations of the FoI. From one end, the FoI would be people, and for another, it would be the city. How would the SIF help a user to compare two different FoIs to each other? and to which extent this can still be mapped to an application ontology with a consistent model without causing a conflict?

Another example of FoI would be the environment around the sensor, such as the existence of some chemical substance in the air, only molecules that hit the sniffer are sensed. The sniffer might not be able to detect it if the source is a meter away, or in a windy environment where the wind goes in the other direction. This can affect the volumetric properties of the sample, such as the density of the sample. Therefore, in experiments where FoI is the sample (most of the CBRN sensors), we should know the algorithm applied to the sample of volumetric properties as they are a function of the volume of the sample.

5

REVIEW OF CURRENT SIF SEMANTIC MODEL

——

5 REVIEW OF CURRENT SIF SEMANTIC MODEL

The SIF Semantic Model is being developed using a convergent approach. Top-down techniques are used to establish the basic structure and to define the core concepts. These are derived from standards such as the SSN Ontology and the OGC SWE Common data model. At the same time, bottom-up techniques fill in the details using different data models, based on standards and specifications for existing deployed sensor systems. These data models are selected to represent as many different sensor systems as possible.

Concepts that are shared across two or more data models are identified, normalized, then integrated into the top-down framework. The result is a graph representing a semantic tree, with the most general concepts at the root and the most specific concepts out on the branches.

The current draft of the SIF Semantic Model consists of six integrated models:

1. W3C Sensor Semantic Network Ontology (SSN)

2. OGC Sensor Web Enablement (SWE)

3. Sensor Integration Framework

4. Integrated Sensor Architecture (ISA)

5. National Imagery Transmission Format (NITF)

6. Motion Imagery Standards Board (MISB) standards

7. Mappings

5.1. Semantic Sensor Network

In the SSN Ontology, the following are described:

1. Sensors

2. Observations

3. Procedures

4. Features of interest

5. Samples

6. Observed properties

7. Actuators

SSN follows a horizontal and vertical modularization architecture. The elementary classes and properties (Sensor, Observation, Sample, and Actuator) of SSN are based on a lightweight and self-contained core ontology called SOSA. SSN and SOSA, together, are able to cover a wide range of applications and use cases, including satellite imagery, large-scale scientific monitoring, industrial and household infrastructures, social sensing, citizen science, observation-driven ontology engineering, and the Web of Things. The main classes of the SOSA/SSN Ontology are listed in Table 1.

Table 1 — SOSA/SSN Ontology Core Classes

CLASS	DEFINITION	EXAMPLE	ONTOLOGY
sosa:Sensor	Device, agent (including humans), or software (simulation) involved in, or implementing, a Procedure. It can be hosted by Platforms, and it responds to a Stimulus or Input data.	A temperature sensor	SOSA
sosa:Actuation	Uses an Actuator to change the state of the world.	The activity of automatically closing a window if the temperature in a room drops below 20 degrees Celsius.	SOSA
sosa:Actuator	A device that is used by, or implements, an (Actuation) Procedure that changes the state of the world.	A window actuator for automatic window control, i.e., opening or closing the window.	SOSA
sosa:Actuatable Property	An actuatable quality (property, characteristic) of a FeatureOfInterest, e.g. for an automatic window control Actuator, the ability of the window to be opened and closed is its ActuatableProperty.	In the example of automatic window control, the ability of the window to be opened and closed is its ActuatableProperty.	SOSA
sosa:Feature OfInterest	The thing whose property is being estimated or calculated in the course of an Observation to arrive at a Result, or whose property is being manipulated by an Actuator, or which is being sampled or transformed in an act of Sampling.	When measuring the temperature of a room, "room" would be the FeatureOfInterest.	SOSA
sosa:Observable Property	An observable quality (property, characteristic) of a FeatureOfInterest, e.g. the temperature of a room.	When measuring the temperature of a room, "temperature" is the ObservableProperty.	SOSA
sosa: Observation	Act of carrying out an (Observation) Procedure to estimate or calculate a value of a property of a FeatureOfInterest.	The temperature value of the room.	SOSA
sosa:Platform	An entity that hosts other entities, particularly Sensors, Actuators, Samplers, and other Platforms.	A satellite, cell-phone, human, animal or etc. may act as Platforms for Sensors or Actuators.	SOSA
sosa:Procedure	A workflow, protocol, plan, algorithm, or computational method specifying how to	The measured temperature of a room differs depending on	SOSA

CLASS	DEFINITION	EXAMPLE	ONTOLOGY
	make an Observation, create a Sample, or make a change to the state of the world (via an Actuator). A Procedure explains the steps to be carried out to arrive at reproducible Results.	where the Sensor is deployed, e.g., distance to the window or heating/cooling sources. The definition of Sensor placement and other possible considerations are defined by the Procedure.	
sosa:Result	The Result of an Observation, Actuation, or act of Sampling.	The value 25 as the temperature of a certain room together with the unit, e.g., Degree Celsius.	SOSA
sosa:Sample	Device, agent (including humans), or software (simulation) involved in, or implementing, a Procedure. It can be hosted by Platforms, and it responses to a Stimulus or Input data.Feature which is intended to be representative of a FeatureOfInterest on which Observations may be made.		SOSA
sosa:Sampler	A device that is used by, or implements, a (Sampling) Procedure to create or transform one or more samples.		SOSA
sosa:Sampling	An act of Sampling carries out a (Sampling) Procedure to create or transform one or more Samples.		SOSA
ssn:Deployment	Describes the Deployment of one or more Systems for a particular purpose. Deployment may be done on a Platform.	A temperature Sensor deployed on a wall, or a whole network of Sensors deployed for an Observation campaign.	SSN
ssn:Input	Any information that is provided to a Procedure for its use.		SSN
ssn:Output	Any information that is reported from a Procedure.		SSN
ssn:Property	A quality of an entity. An aspect of an entity that is intrinsic to and cannot exist without the entity.		SSN
ssn:Stimulus	An event in the real world that 'triggers' the Sensor. The properties associated to the Stimulus may be different to the eventual observed ObservableProperty. It is the event, not the object, that triggers the Sensor.		SSN
ssn:System	System is a unit of abstraction for pieces of infrastructure that implement Procedures. A System may have components, its subsystems, which are other Systems.		SSN

5.2. Sensor Web Enablement

The OGC's SWE standards enable developers to make all types of sensors, transducers, and sensor data repositories discoverable, accessible and useable via the Web. It allows applications and/or servers to structure, encode and transmit sensor datasets in a self-describing and semantically enabled way.

SWE Common is a semantics-free syntax (uses identifiers pointing to external semantics) that is part of the SWE suite of standards. A key concept of the SWE Common Data Model is the ability to separate data values themselves from the description of the data structure, semantics, and representation. A SWE dataset description can be connected to an existing taxonomy encoded in a Web Ontology Language (OWL) document, or a Geography Markup Language (GML) dictionary, that provides a unique identifier for each entry.

1. Data representation types can be Boolean, Categorical, Continuous Numerical, Discrete Countable, or Textual.

2. The Nature of Data is connected to external semantic resources such as dictionaries, taxonomies, or ontologies.

3. Data encoding methods define how the data is packed as blocks (Packages) that can efficiently be transferred or stored using various protocols and formats

In a Sensor Model Language (SensorML) document that has a `definition` attribute, the value of that attribute should resolve to a definition provided by the OGC Definitions Server, which gives the meaning of that particular element. There are so many things that are general classifications, such as parameter, property, configuration element, etc. Equivalence map matching of a raw property to the semantic equivalent of another source and presenting it to the user as a single integrated semantic framework is so valuable. This mapping requires knowledge on transforming from one ontology to another.

5.3. Sensor Integration Framework

The SIF Semantic Model is where all of the integration between the higher-level SSN and SWE models and the lower-level domain models takes place. It's the result of a convergence process, where higher-level concepts are specialized to the point where they are mappable to domain concepts.

5.4. Integrated Sensor Architecture

The Integrated Sensor Architecture (ISA) is a framework developed by the U.S. Army for integrating tactical sensors. Its focus is primarily on ground sensors operating in Denied, Disrupted, Intermittent, and Limited Bandwidth (DDIL) networks. The technical limits faced by these systems made them an ideal use case for the development of the SIF. As a result, the ISA data model is well integrated into the SIF Semantic Model.

The ISA Model is captured in the SIF Semantic Unified Modeling Language (UML) model. In addition, an ontology of ISA properties was generated in the course of this initiative and is available at http://sensorml.com/ont/isa/property. This is the most mature Technical View in the SIF Semantic Model.

5.5. National Imagery Transmission Format

The National Imagery Transmission Format (NITF) is a standard developed for the U.S. Department of Defense and Intelligence Community. It is a container format, primarily designed for still imagery, but flexible enough to accommodate other payloads as well. The ISO 12087 Basic Image Interchange Format (BIIF) is based on NITF.

The NITF format is of particular interest due to the rich metadata which it can carry. This is commonly described as exploitation metadata, everything you need to know to extract information out of an image.

Due to the size of NITF, the SIF Semantic Model focused on those parts most likely to be mappable to other sensor models. Those include:

1. Information about the sensor system and supporting platform

2. Information about the collection mission

3. General information about the image and its structure

4. Photogrammetric information

5. Image formation information

This metadata is represented in the SIF Semantic UML model. The current version only captures the elements and their definitions. Note that a previous OGC Testbed examined NITF in relation to GMLJP2 (OGC 12-154).

The NITF Technical Board (NTB), the organization responsible for the NITF Standard, has developed a draft Motion-Imagery-Derived Still Imagery (MIDSI) standard. This standard will define how to populate NITF metadata for a single frame from a MISB stream. It includes

detailed transformation logic to convert MISB metadata elements into NITF metadata elements. This information will be exploited for future versions of the SIF Semantic Model.

5.6. Motion Imagery Standards Board

The Motion Imagery Standards Board (MISB) is responsible for developing motion imagery standards for the U.S. Department of Defense and Intelligence Community. Their standards extend the commercial motion imagery standards developed by the Society of Motion Picture and Television Engineers (SMPTE).

The MISB Standard 0601 Unmanned Air System (UAS) Datalink Local Set standard defines the top-level metadata delivered as part of a motion imagery data stream. MISB Standard 0601 is the core metadata standard for motion imagery. This metadata is time synchronized with the motion imagery to provide detailed telemetry about each frame. The primary purpose of this metadata is to enable the extraction of ground coordinates from the motion imagery.

This MISB metadata is represented in the SIF Semantic UML model. It is both incomplete and insufficient for the job. In the course of Testbed 17, an effort was made to mature the MISB model for those elements defined in MISB Standard 0601. That effort and the lessons learned are described in the Clause 6.11 section of this report.

5.7. Mappings

The Mapping section defines how elements in a Domain model map into SIF elements. The current approach uses `mapsto` associations between UML classes. The source of the association is the SIF class while the target is a property of a Domain class.

This approach can only work if the SIF classes are carefully selected. The following section describes how multiple forms of rotation have been represented in the SIF Semantic Model. This representation has proven effective for mapping rotation elements from a number of Domain models.

The rotation class is an abstraction of four different ways of Implementing rotations. All four of these classes are the same concept (rotation) but the mathematics used to perform the rotation differ. The domain elements are specific to one or more mathematical techniques. But the conceptual model should be independent of technique.

1. Establish each of the technique-specific classes as specializations of Rotation.

2. Populate each technique-specific class with the properties needed to implement that mathematical technique

3. Map the elements of the domain model into the properties

4. Define operations on the rotation class to convert between each technique-specific class pair.
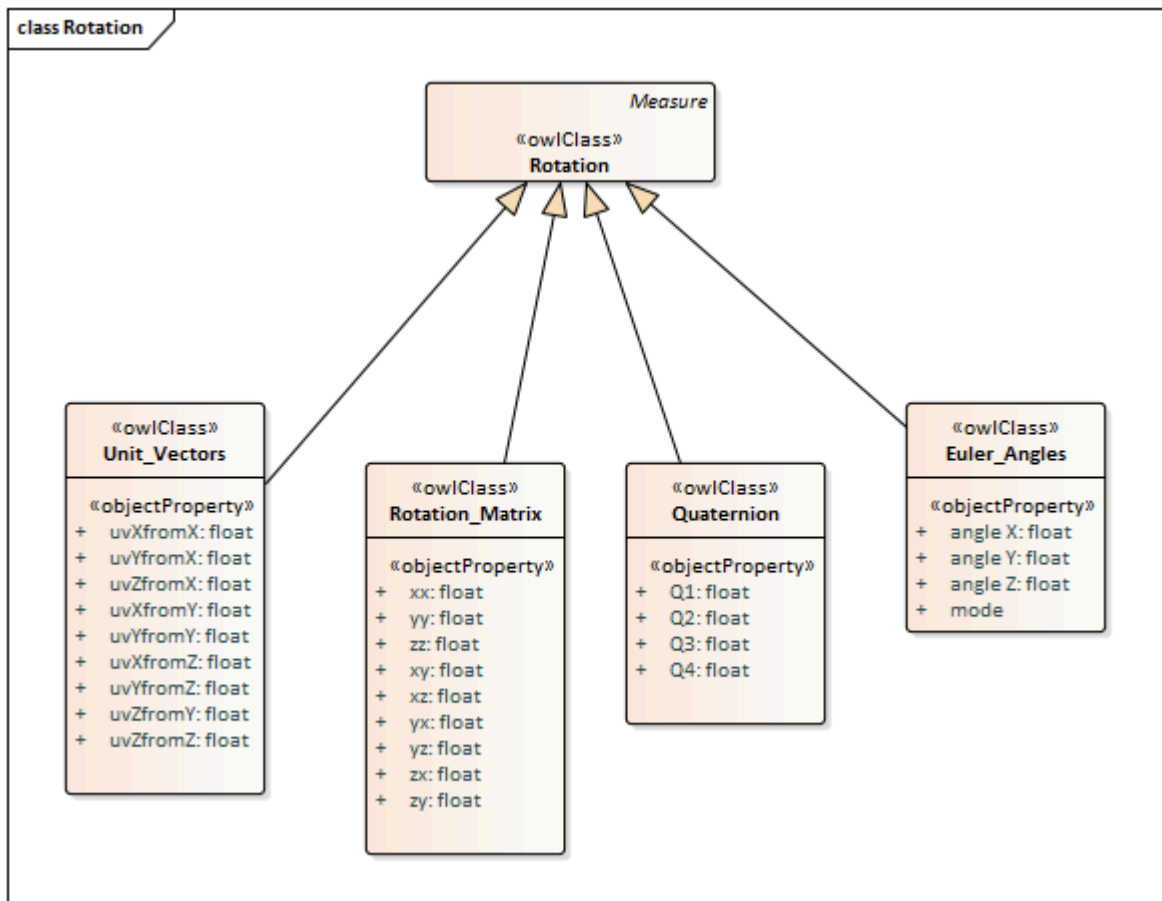
Figure 1 — SIF rotation class

A similar approach can be used for location elements. This model has worked well for two and three-dimensional systems. Additions will be needed when we encounter n-dimensional and non-euclidean spaces.
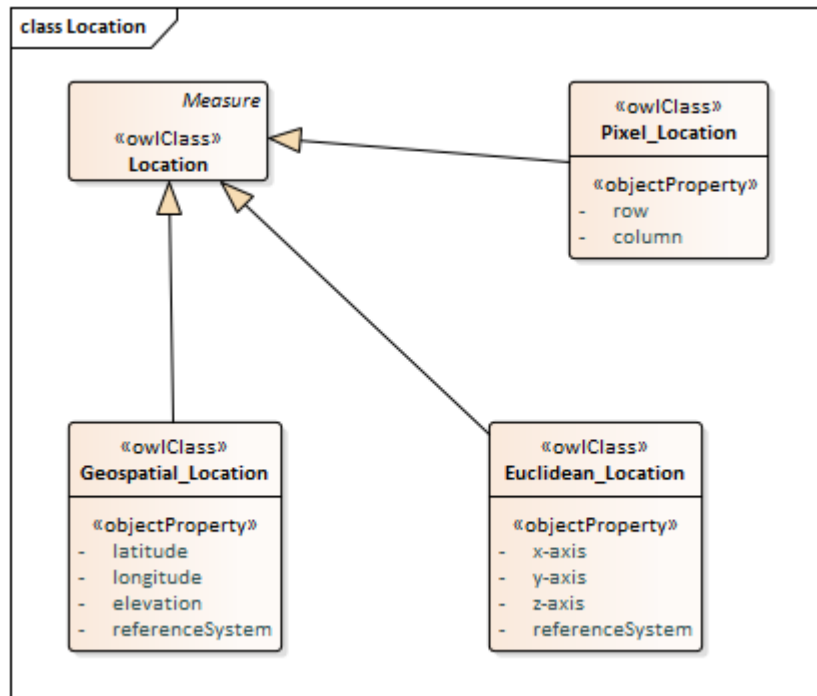
Figure 2 — SIF location class

These approaches can work because we are dealing with mathematics. There is no ambiguity in how A converts to B because the math is well established.

6

IMPLEMENTING THE SEMANTIC MODEL

6 IMPLEMENTING THE SEMANTIC MODEL

One of the goals of the SIF work item was to evaluate if and how effectively the SIF Semantic Model performed its job. This evaluation involved taking three domain models (ISA, NITF, and MISB), implementing them using appropriate technologies, and then evaluating their ability to convey semantically complete information.

From this exercise we learned that the SIF Semantic Model has a long way to go. But a number of valuable lessons were learned and practices developed.

6.1. UML

The SIF Semantic Model has been developed using the Sparx Systems Enterprise Architect product and UML. One could argue that UML is not an appropriate technology to represent a semantic model. That argument has some validity. However, the first task in constructing a semantic model is knowledge capture. UML has proven to be an effective tool for knowledge capture.

1. Most domain experts have a basic understanding of UML and can discuss concepts represented using UML diagrams. We have not found this to be true of other tools.

2. Many of the data models that are being examined already have at least some UML-based documentation.

3. UML captures concepts, associations, and operations. The last is important to capture transformation logic.

While there was general agreement that a more semantics-native approach is desirable, no such approach has been identified. Nor is it likely that such an approach can be identified until the other issues identified by this Testbed effort (Clause 7) are resolved. Those resolutions comprising the requirements that a semantic modeling technique must address.

6.2. Ontologies

Many of the entities defined by the SWE Common Standard include a `definition` attribute. This attribute allows the entity to reference its semantic definition as expressed through an ontology. Implementations of the SIF make use of this feature to separate semantic and syntactic translation. Therefore, the SIF Semantic Model must provide OWL/RDF ontologies in addition to the UML.

In order to provide a common representation of SIF semantics, the following conventions were applied:

- All SIF ontologies are hosted on the <u>http://sensorml.com/ont/</u> root. For example, the MISB Standard 0601 ontology is at <u>http://sensorml.com/ont/misb0601</u>.

- SIF concepts are instantiated as `owl:NamedIndividual` objects.

- All `owl:NamedIndividual` objects include:

  - `skos:definition` which contains a human readable definition of the object.

  - `rdfs:label` which contains a short name for the object.

  - `rdfs:comment` which contains the name of the governing standard and (if appropriate) section number where this element is defined.

  - `skos:broadMatch` used to identify concepts which this object specializes. This association is used to identify upper-level concepts from the SSN, SWE, or SIF ontologies.

  - `ssn:isPropertyOf` identifies the associated Feature of Interest.

- In addition, the following properties may appear in a SIF `owl:NamedIndividual` object:

  - `sif:authority` identifies the `authority` which defines a qualified name or value.

  - `skos:broadMatch` identifies concepts from external ontologies which are specialized by this element.

  - `nsl:applicableUnit` identifies the Unit of Measure.

  - `sif:referenceSystem` identifies the coordinate reference system used

  - `skos:exactMatch` identifies two objects that are different representations of the same concept.

This list of conventions is at best a preliminary set.

6.3. Generating Ontologies

Ontology generation begins with a set of conventions for the UML models. These conventions seek to align, as much as possible, the meta-model of UML with the meta-model of OWL.

Those conventions are as follows:

- Each data element is captured as a UML class

- With the exception of enumerations, the UML classes have the `owl:class` stereotype.

- With the exception of subclass relationships, all associations are captured as attributes of the class.

- These attributes have the `owl:objectProperty` stereotype.

- With the exception of subclass relationships, all associations are captured as UML associations between the relevant classes.

- All UML associations are directional (source to target)

- The role of the target of an association is selected from the attributes of the target.

- UML generalization associations represent subclass associations.

The result is a conceptual model which is complete, easy to navigate, and easy to validate.

The next step is to generate the OWL ontology from the UML model. This is easier said than done.

Enterprise Architect has a feature to publish a UML model as OWL/RDF. The results are not very useful.

```
<owl:Class rdf:ID="Image_Coordinate_System">
    <rdfs:comment>Name of the image coordinate system used </rdfs:comment>
</owl:Class>
```

Figure 3 — Code Example of OWL/RDF exported from Enterprise Architecture

The Enterprise Architect tool allows users to build document templates and then use those templates to generate Rich Text Format (RTF) representations of the model. This technique was used to generate AsciiDoc content for the CityGML 3.0 Standard. It stands to reason that this technique should also work for generating an OWL/RDF content.

Note: class names cannot contain spaces.

```
package >
element >
    <owl:NamedIndividual rdf:about="http://sensorml.com/ont/misb0601/{Element.
Name}">
        <skos:definition>
          {Element.Notes}
        </skos:definition>
        <skos:broadMatch rdf:resource="http://sensorml.com/ont/swe/property/tbd"/
>
        <skos:broadMatch rdf:resource= "http://qudt.org/vocab/quantityKind/tbd"/>
        <ssn:isPropertyOf rdf:resource="http://sensorml.com/ont/misb0601/object"/
>
        <nsl:applicableUnit rdf:resource="http:qudt.org/vocab/unit/M/"/>
        <referenceSystem rdf:resource="https://apps.epsg.org/api/v1/Datum/5100"/>
        <authority rdf:resource="http:www.example.com"/>
        <rdfs:label>{Element.Name}</rdfs:label>
        <rdfs:comment "See MISB 0601.15 section TBD"/>
```

```
    </owl:NamedIndividual>
  < element
  < package


  <owl:NamedIndividual rdf:about="http://sensorml.com/ont/misb0601/Image_
Coordinate_System">
      <skos:definition>
        Name of the image coordinate system used
      </skos:definition>
      <skos:broadMatch rdf:resource="http://sensorml.com/ont/swe/property/tbd"/
>
      <skos:broadMatch rdf:resource= "http://qudt.org/vocab/quantityKind/tbd"/>
      <ssn:isPropertyOf rdf:resource="http://sensorml.com/ont/misb0601/object"/
>
      <nsl:applicableUnit rdf:resource="http:qudt.org/vocab/unit/M/"/>
      <referenceSystem rdf:resource="https://apps.epsg.org/api/v1/Datum/5100"/>
      <authority rdf:resource="http:www.example.com"/>
      <rdfs:label>Image_Coordinate_System</rdfs:label>
      <rdfs:comment "See MISB 0601.15 section TBD"/>
    </owl:NamedIndividual>
```

Figure 4 — Code Example of RTF representations of the model

The RDF template shown above implements the following rules. These rules are far from complete.

- UML Classes are instantiated as `owl:NamedIndividual` objects.

- For all `owl:NamedIndividual` objects:

  - `skos:definition` property is populated from the UML class description.

  - `rdfs:label` property is populated with the name of the UML class.

  - `rdfs:comment` property is populated with the name of the standard and (if appropriate) section number where this element is defined. This element requires manual editing.

  - `skos:broadMatch` identifies concepts which this object specializes. This association is used to identify upper-level concepts from the SSN, SWE, or SIF ontologies. Currently this requires manual editing.

  - `ssn:isPropertyOf` identifies the Feature of Interest for this property. Currently this requires manual editing.

6.4. Features of Interest

The data models for sensor systems often do not define the Features of Interest that are the target of their sensors. So, they have to be derived from the supporting narrative of the data model. This is also true for the components of a sensor system. Measures describing the location, orientation, and state of a component are properties of that component. So, a sensor system component also has to be treated as a Feature of Interest.

A feature of interest is commonly viewed as a tangible "thing" that exist in the real world. For a sensor system, those "things" may be ephemeral. The relative wind measure, for example, a measure of the direction of the wind relative to a flying platform. The feature of interest may be the atmosphere (wind) in the immediate vicinity of the platform. A feature of interest that is continuously changing may also be observed by a sensor.

Some progress was made in this Testbed on defining these more ephemeral features of interest. But a broader consensus is needed if the community is to build on this work.

6.5. Data Types

Sensor properties may be used in mathematical computations. The mathematical type of the property is an important part of its identity. Rather than create a Type taxonomy, the quantityKind ontology at http://Qudt.org/vocab/quantityKind was used.

The `Skos:broadMatch` association was used to indicate that the sensor parameter is a specialization of the referenced quantityKind

Not all data types are covered by QUDTs. The OntoDT ontology was leveraged to fill the gaps.

6.6. Units of Measure

The QUDT `qualityKind` ontology identifies a range of Unit of Measure (UoM) definitions for each concept. But the SIF needs to be able to specify a single UoM. Mixing UoM systems in a mathematical computation can lead to invalid results.

The QUDT Units ontology was selected as the source for most units of measure. The QUDT ontology is not comprehensive, so an alternative was needed for those UoM which are not supported by QUDT.

The NSG Physical Quantities Vocabulary Register was able to fill some of the gaps. Additional registers will be identified as needed.

6.7. Coordinate reference systems

Sensor data is different from other geospatial data in that the actual geolocation is often not relevant. Most calculations are performed in engineering coordinate reference systems. Geo-coordinates are only applied once all of the engineering calculations have been performed.

Most sensor systems do not have a single engineering coordinate reference system. Each component in the system may have its own CRS. The value of all measures reported from that component are in the local CRS. A SIF model for a sensor system must identify every CRS used

in the system as well as the rules for transforming values between those CRS. This is a formable task when given a complex, multi-articulated sensor system.

MISB Standard 1906 was selected to address this issue. MISB Standard 1906 describes the Motion Imagery Metadata Staging System. This is a model for spatial relations between the components of a sensing system. Each component of the system is a `stage`. Stages can have relative and geospatial locations. Relative locations between stages are defined by the translation (offsets), rotation, and kinematics (velocity, acceleration, etc.). Geospatial locations provide an `anchor` to a geospatial reference systems (usually WGS84E-3D).

MISB Standard 1906 may provide the foundation of a more complete understanding of engineering CRS geometries. Such a system of geometry is essential for further integration of sensor systems as well as other non-geo spatial systems.

6.8. Temporal Reference Systems

Many properties of a sensor system are only valid for a limited (sometimes quite short) time. A temporal model is needed which can represent the time sensitivity of the metadata values.

Further complicating matters, a sensor system may have multiple clocks which may or may not be synchronized. So, the the concept of a "Timer" was borrowed from MISB Standard 1603.

A `Timer` serves the role of a coordinate reference system for time values when precision and accuracy are important. In addition, relationships between Timers are captured allowing a measurement from one Timer to be converted into an equivalent measurement from another Timer.

6.9. Domain to SIF Transformation

In order to enable software-based transformation of information from a Domain model into a normalized SIF model, the transformation rules for each element must be defined. Ideally they would be included as part of the SIF semantic framework. This proved to be much harder than anticipated.

The simplest case is whereby the concepts are equivalent. The `skos:exactMatch` association addresses this case quite well. Unfortunately this case is not as common as we would like.

A second case is where the SIF concept has a higher level of abstraction than the domain concept. This looks to be the most common case. The `skos:broadMatch` association captures this concept.

However, our goal is to represent sufficient information that both syntax and semantic transformation can be performed by software.

One option could be the use of a formal pseudocode language to capture transformation logic. This is similar to the approach used in the MISB MIDSI draft standard. This standard defines how to generate a NITF still imagery file from a MISB frame. The primary focus of this standard is the conversion of MISB metadata into NITF metadata. Extensive use of pseudocode is employed for this purpose.

This approach could be adapted for the SIF through use of the `Simple Process` object from SensorML. `Simple Process` would serve as a container for the process logic needed to transform specific inputs from the domain model into SIF outputs. The process logic itself would be defined in the associated `Process Method` object. That logic, in turn, could use standardized algorithms (abstract algorithm object) to build more complex workflows.

At this point this content is only a hypothetical. Additional research is needed to validate that this approach would actually work and will scale.

6.10. The NITF Ontology

The only ontology generated for NITF was created for the SENSRB Tagged Record Extension (TRE). Since SENSRB covers sensor parameters for airborne imaging Electro-Optical sensors, it is a good compliment to the MISB work.

Unfortunately, generation of an ontology to the level of detail as the MISB Standard 0601 ontology was not feasible in this Testbed. The biggest issue is that some of the SENSRB terms are polymorphic. They take on different meaning based on the value of other parameters. A complete ontology would have to identify all of these polymorphic entities and derive all possible sub-concepts. A task that could not be accomplished within the budget and schedule of this initiative.

As a result, the SENSRB ontology is a dictionary of terms and definitions, generated from the UML model.

6.11. The MISB Ontologies

MISB metadata standards define how to encode metadata related to the collection of Motion Imagery using the Key-Length-Value (KLV) encoding standard (SMPTE ST 336) developed by SMPTE. This metadata is delivered along with, and time-synchronized with the motion imagery stream. It is optimized for real-time streaming while consuming minimum bandwidth.

The goal of this effort was to build an ontology that fully captures the concepts documented in MISB Standard 0601. That requires not just capturing discrete concepts, but also capturing how those concepts fit together into a mathematically coherent representation of the sensor system.

The first step was to build a UML model capturing the content of MISB Standard 0601. That in itself was not too difficult since the MISB documentation is clear and fully describes each metadata element. The process was as follows:

1. Created one class for each tag in the MISB Standard,

2. Set the stereotype for each class as `owlClass`,

3. Populated the description of each class with the description from the MISB Standard.

At this point, it was a simple step to export the UML model out of Enterprise Architect in an RDF/OWL representation.

However, the resulting ontology was far from acceptable. There were a number of issues and it was not clear how they could be addressed using UML. So, from that point on, most developments were performed by hand-editing the RDF/OWL document.

7

ISSUES AND DISCUSSIONS

7 ISSUES AND DISCUSSIONS

The SIF Semantic Model was developed by the SIF Working Group. The SIF Working Group is a subgroup of the Geospatial Web Services (GWS) Focus Group of the Geospatial-Intelligence Standards Working Group (GWG). Participants in its development came almost exclusively from the U.S. Defense and Intelligence community.

The SIF was brought to the OGC with the goal of receiving comments from a much larger body of reviewers. The following sections discuss the issues and recommendations identified through this review.

7.1. Issue: Scope of the Semantic Model

ISSUE   The SIF Semantic Model is designed to provide run-time translation between concepts and formats. Is that goal realistic or should a less ambitious goal be established?

The ultimate goal of the SIF effort was to provide complete transparency across sensor systems. Users can access and exploit both sensors and observations regardless of the underlying information models. This implies that sensors and observations are transformed in real-time into a normative representation. Such a transformation requires that the semantic model represents not just concepts, but also the transformation rules for converting between similar concepts.

Experience in implementing the SIF led to moving the target of normalization away from the Reference model to the Enterprise mode. That is to say, the SWE standards form the normalized model upon which all domain models are mapped.

1. It is already in many ways a platform-independent model

2. The underlying encoding standard, SWE common, separates syntax from semantics.

Early in this initiative, the testbed participants were able to demonstrate syntactic translation using OpenSensorHub (OSH) and MASBUS as bridging services between the domain-specific models and common SWE model.

These services did not demonstrate semantic translation. However, by capturing domain semantics in a publicly accessible register (http://www.sensorml.com), they could preserve their domain semantics through the syntactic conversion.

This validates the approach for associating semantics and syntax. But does not address the syntax transformation issue.

7.2. Issue: Depth of the Model

ISSUE   The SIF Semantic Model was developed under the philosophy that any concept that could be mapped should be mapped. Is that expansive approach justified, or should a more focused approach be used?

The SIF Semantic Model was developed without any restrictions on the conceptual elements that would be considered. This approach was taken for two reasons. First, the developers were concerned that the model would not be scalable. By including the hard cases, the resulting model would work for the easier cases as well. The second reason was the difficulty of selecting concepts to include. Software analytics is expected to be a major consumer of SIF data. Identifying the sensor concepts which will be meaningful to software analytics is a daunting if not impossible task. Limiting the effort to a pre-selected set of concepts would result in a flawed model which would have limited use.

Working with the MISB and ISA models revealed an underlying order to these large data models. The concepts sort themselves into a smaller number of common patterns. These patterns appear to align neatly with the classes defined in SWE Common. A result is a limited number of classes that are instantiated in an unlimited number of named instances. Such a model is much easier to manage since, while the model is large, it consists of a small number of common classes with known behavior and information content.

7.3. Issue: Use of UML

ISSUE   The decision to use UML for the Semantic Model was made a couple of years ago and reflects a limited knowledge of the available tools. Are these conclusions still valid?

The SIF Semantic Model has been developed using the Enterprise Architect tool and UML. One could argue that UML is not an appropriate technology to represent a semantic model. That argument has some validity. However, the first task is knowledge capture, not knowledge representation. UML has proven to be an effective tool for knowledge capture.

1. Most domain experts have a basic understanding of UML and can discuss concepts represented using UML diagrams. We have not found this to be true of other tools.

2. Many of the data models that are being examined already have at least some UML-based documentation.

3. UML captures concepts, associations, and operations. The last is important to capture transformation logic.

7.3.1. MISB Ontology

The MISB Standard 0601 was chosen as the starting point for a motion imagery metadata ontology since all other MISB metadata standards are extensions to this data model.

MISB Local Sets define motion imagery metadata concepts and how they should be encoded using SMPTE, KLV format. KLV encodes each metadata element as an atomic value. Each value is reported separately. As a result, the metadata elements have no explicit associations. For example, it is legal to report the latitude but not the longitude of a platform.

Since the associations between metadata elements are not explicit in the Standards, a UML model proved valuable as a tool to capture the individual elements and derive the associations between them.

The following conventions were used when capturing the MISB domain knowledge:

- Each metadata element is captured as a UML class

- With the exception of enumerations, the UML classes have the `owl:class` stereotype.

- With the exception of subclass relationships, all associations are captured as attributes of the class.

- These attributes have the `owl:objectProperty` stereotype.

- With the exception of subclass relationships, all associations are captured as UML associations between the relevant classes.

- All UML associations are directional (source to target)

- The role of the target of an association is selected from the attributes of the target.

- UML generalization associations represent subclass associations.

The result was a conceptual model of MISB Standard 0601 which is easy to navigate, complete, and easy to validate.

The next step was to simply generate the owl ontology from the UML model. Easier said than done.

Use of RDF export

Enterprise Architect has a feature for publishing a UML model as OWL/RDF. See Figure 3 for an example of OWL/RDF code exported from Enterprise Architect. The testbed participants did not find the results useful for representing the SIF.

Use of comprehensive template

The Enterprise Architect tool allows users to build document templates and generate RTF representations of the model using those templates. This technique was used to generate AsciiDoc content for the CityGML 3.0 Standard. It stands to reason that this technique should

also work for generating an OWL/RDF content. An example of RTF representations of the model is shown in Figure 4. Note that class names cannot contain spaces.

Use of minimal template

In order to provide a common representation of SIF semantics, a template with the following conventions was applied:

- All SIF ontologies are hosted on the <u>http://sensorml.com/ont/</u> root. For example, the MISB 0601 ontology is at <u>http://sensorml.com/ont/misb0601</u>.

- UML Classes are instantiated as `owl:NamedIndividual` objects

- All `owl:NamedIndividual` objects include:

  - `skos:definition` properties populated from the UML class description

  - `rdfs:label` properties populated with the name of the UML class.

  - `rdfs:comment` properties populated with the name of the standard and (if appropriate) section number where this element is defined. This element requires manual editing.

  - `skos:broadMatch` identifies concepts which this object specializes. This association is used to identify upper-level concepts from the SSN, SWE, or SIF ontologies.

  - `ssn:isPropertyOf` identifies the Feature of Interest for this property

Identifier concepts

- sif:authority

Quantity concepts

- `skos:broadMatch` identifies concepts from external ontologies which are specialized by this element. In most cases this is the <u>http://qudt.org/vocab/quantityKind</u> ontology.

- `nsl:applicableUnit` the Unit of Measure. Typically from <u>http://qudt.org/vocab/unit</u>.

- `sif:referenceSystem`

The construction of the MISB ontology is described in the Clause 6.11 section of this report.

8

FUTURE WORK

FUTURE WORK

The following future work items have been identified:

8.1. Multi-level Sensor Ontology Architecture

Define a best practice architecture to split the scope and responsibility of different levels of sensor and observable semantics into several ontologies maintained by different organizations, to provide more flexibility when concepts need to be added. At least 3 levels could be defined, such as:

- A cross-domain ontology hosted by OGC and maintained by OGC members (an extension of the work done on W3C/OGC SSN and QUDT)

- Domain-specific ontologies hosted by separate organizations (typically organizations whose responsibility includes standardization over a specific application domain such as the World Meteorological Organization (WMO), North Atlantic Treaty Organization (NATO), etc.)

- Local ontologies defined by independent users with particular needs and typically hosted on the data servers themselves

The idea would be to maintain cross-ontology mappings so that specific concepts (domain-specific or even equipment-specific) can refer to more generic concepts in ontologies located at a higher level in the hierarchy. Simple relationships such as SKOS mapping properties (skos:closeMatch, skos:exactMatch, skos:broadMatch, skos:narrowMatch, skos:relatedMatch) could be used initially.

Another important aspect would be to give visibility to all ontologies that would be part of this semantic framework so they can easily be searched for existing concepts, to avoid a proliferation of duplicated concepts.

Finally, governance aspects should also be addressed so that new additions to the semantic tree can be discussed and potentially moved up in the tree when applicable to more than one domain or deployment.

8.2. Cross-Domain Sensor Ontology

Develop an OGC cross-domain ontology (mentioned above) that includes concepts that are applicable to many sensor types. It could be hosted on the OGC Definitions Server. The exact

scope of the ontology is to be defined but needs to be built on SSN. It could include mappings to corresponding Wikipedia articles and cover the following aspects:

- Common system types taxonomy (extension of SSN)

- Common sensor/actuator types taxonomy

- Common platform types taxonomy

- Common system identifiers (serial number, model number, etc.)

- Common system classifiers (sensor type, application, etc.)

- Common system properties (extension of SSN)

- Common system capabilities (extension of SSN)

- Positioning concepts (heading, bearing, orientation as Euler angles, rotation matrices or quaternions, complete pose (position + orientation, etc.) and their derivatives

- Uncertainty and statistical concepts (accuracy, precision, circular error, probability distributions, etc.)

- Response model characteristics (spectral response, directional response, temporal response, etc.)

8.3. More Formal Mapping Definitions

Ultimately, one may want to formally define the mathematical relationships between some of the concepts. More advanced mapping ontologies could be developed to that end.

A

ANNEX A (INFORMATIVE) REVISION HISTORY

A ANNEX A (INFORMATIVE) REVISION HISTORY

DATE	EDITOR	RELEASE	PRIMARY CLAUSES MODIFIED	DESCRIPTIONS
November 18, 2021	M. Jahromi	.1	all	initial version
November 22, 2021	M. Jahromi	.8	all	complete version
January 06, 2022	M. Jahromi & A. Robin	1.0	all	editorial changes & future work
February 06. 2022	M. Jahromi	1.1	all	editorial changes

BIBLIOGRAPHY

BIBLIOGRAPHY

1.   Simon Cox: OGC 10-004r3, *Topic 20 — Observations and Measurements*. Open Geospatial Consortium (2010). https://portal.ogc.org/files/?artifact_id=41579

2.   Armin Haller, Krzysztof Janowicz, Simon Cox, Danh Le Phuoc, Kerry Taylor, Maxime Lefrançois: OGC 16-079, *Semantic Sensor Network Ontology*. Open Geospatial Consortium (2019). https://www.w3.org/TR/2017/REC-vocab-ssn-20171019/

3.   NGA: MISB ST 0601 UAS Datalink Local Set standard. National Geospatial Intelligence Agency, https://gwg.nga.mil/misb/docs/standards/ST0601.16a.pdf

4.   OMG: Model Driven Architecture. Object Management Group, https://www.omg.org/mda/

5.   NGA: Sensor Integration Framework (SIF). National Geospatial Intelligence Agency, https://github.com/ngageoint/Sensor_Integration_Framework

6.   Darko Androsevic: OGC 12-154, *OWS-9 OWS Innovations GMLJP2 for National Imagery Transmission Format (NITF) Engineering Report* . Open Geospatial Consortium (2013). https://portal.ogc.org/files/?artifact_id=51889

7.   NGA: WGS84E-3D. National Geospatial Intelligence Agency, https://nsgreg-api.nga.mil/coord-ref-system/WGS84E_3D

8.   OGC: Testbed-17 ISA Ontology, http://sensorml.com/ont/isa/property

9.   George Percivall: OGC 13-032, *OGC® SWE Implementation Maturity Engineering Report*. Open Geospatial Consortium (2013). https://portal.ogc.org/files/?artifact_id=53823

10.   Alexandre Robin: OGC 08-094r1, *OGC® SWE Common Data Model Encoding Standard*. Open Geospatial Consortium (2011). https://portal.ogc.org/files/?artifact_id=41157

11.   W3C: Mathematical Markup Language (MathML) Version 3.0 , World Wide Web Consortium, https://www.w3.org/TR/MathML3/ (2014)