

Unifying Universal Core, SUMO, OWL 2, and XML Standards to Build Intelligence Ontologies

Semantic Technology Conference 2011

Richard Lee

lee_richard@bah.com

Booz | Allen | Hamilton

Overview

- Project 1: METS
- Project 2: SCSP
- Examples
- Future Work

Metadata Extraction and Tagging Service (METS)

- An enterprise service for processing documents:
 - Normalization
 - Metadata detection
 - Entity extraction
 - Geotagging
- Produces XML:
 - An Intelligence Community (IC) standard called PUBS
 - OWL/RDF

Metadata Extraction and Tagging Service (METS)

- Metadata (think Dublin Core, extended):
 - Security markings
 - Author, Publisher, PoC, etc
 - Title, Subtitle, Description, Summary, Dates, etc
 - Format, MIME type, Size, etc
- Extraction results:
 - Events, Persons, Organizations, Locations, Equipment, Dates, Phone Numbers, etc
 - Properties and Relations

Ontology Work on METS

- Industry (OWL and KIF) Standards and near-Standards:
 - SUMO / MILO / etc (converted from KIF)
 - W3C's GML (supplemented) and Time
 - ISO, FIPS, et al code lists converted to OWL enumeration classes
- Government (XML) Standards:
 - TWPDES (for person data)
 - ISM (for security markings)
 - DDMS (for other metadata)

SCSP

- Analysts needed to work with:
 - multiple data sources
 - multiple process models
 - multiple analysis models
- Analysts needed to perform:
 - search
 - discovery
 - correlation
 - presentation

SCSP

- Building ontologies by converting models
 - Palantir
 - MIDB (partial), TIDE, Artemis, et al
 - CIA World Fact Book, ProMED, et al
 - PMESII, CTAF
 - various process models
- Building a “master” ontology
 - Defines all the concepts of interest to the analysts
 - Defines the relationship between the concepts in all the others

Example 1: TWPDES supplementing SUMO

```
<owl:Class rdf:ID="Company">  
  <rdfs:subClassOf rdf:resource="#Organization"/>  
  <owl:equivalentClass  
rdf:resource="&sumo;Corporation"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="Government">  
  <rdfs:subClassOf rdf:resource="#Organization"/>  
  <owl:equivalentClass  
rdf:resource="&sumo;GovernmentOrganization"/>  
</owl:Class>
```

```
<owl:Class rdf:ID="NGO">  
  <rdfs:subClassOf rdf:resource="#Organization"/>  
  <!-- SUMO does not appear to have a match -->  
  <rdfs:subClassOf  
rdf:resource="&sumo;NonprofitOrganization"/>  
  <rdfs:disjointWith rdf:resource="#Government"/>  
</owl:Class>
```

```
<owl:ObjectProperty rdf:ID="affiliation">  
  <rdfs:comment>Indicates some sort of membership  
or affinity from the Person to the  
Organization</rdfs:comment>  
  <rdfs:subPropertyOf  
rdf:resource="&sumo;associateInOrganization"/>  
  <rdfs:domain rdf:resource="#Person"/>  
  <rdfs:range rdf:resource="#Organization"/>  
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:ID="memberOf">  
  <rdfs:comment>Indicates the Person is a member  
of the Organization</rdfs:comment>  
  <rdfs:subPropertyOf rdf:resource="#affiliation"/>  
  <owl:equivalentProperty  
rdf:resource="&sumo;member"/>  
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:ID="employedBy">  
  <rdfs:comment>Indicates the Person is an  
employee of the Organization</rdfs:comment>  
  <rdfs:subPropertyOf rdf:resource="#affiliation"/>  
  <owl:inverseOf rdf:resource="&sumo;employs"/>  
</owl:ObjectProperty>
```


Example 2 Problem

- Equipment vs Sensor vs Vehicle vs Weapon vs etc
 - Artemis: Communication (Device), Equipment are distinct
 - MEPED: Equipment types include Communication Device, Sensor, Vehicle and Weapon, but only covers Military Equipment
 - Palantir: Equipment has children Communication Device, Sensor, et al, but Vehicle and Weapon are siblings of Equipment
 - UCore SL: Sensor vs Equipment disjoint; Vehicle vs Equipment unspecified; Weapon "in principle" would be a child of Equipment
 - (bio) DB: Equipment even includes the Kitchen Sink, and Lab Animals

Example 2 Solution: Interpose SUMO's Device

```
<owl:Class rdf:ID="Device">
  <rdfs:subClassOf rdf:resource="#Artifact"/>
  <owl:disjointWith rdf:resource="#Facility"/>
  <rdfs:subClassOf rdf:resource="&sumo;Device"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Equipment">
  <rdfs:subClassOf rdf:resource="#Device"/>
  <owl:disjointWith rdf:resource="#ExplosiveDevice"/>
  <owl:disjointWith rdf:resource="#Sensor"/>
  <owl:disjointWith rdf:resource="#Vehicle"/>
  <owl:disjointWith rdf:resource="#Weapon"/>
  <rdfs:subClassOf rdf:resource="&art;Equipment"/>
</owl:Class>
```

```
<owl:Class rdf:ID="MeasuringDevice">
  <rdfs:subClassOf rdf:resource="#Device"/>
  <owl:disjointWith rdf:resource="#Communi.Device"/>
  <owl:disjointWith rdf:resource="#ExplosiveDevice"/>
  <owl:disjointWith rdf:resource="#Vehicle"/>
  <owl:disjointWith rdf:resource="#Weapon"/>
  <rdfs:subClassOf rdf:resource="&sumo;M.Device"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Sensor">
  <rdfs:subClassOf rdf:resource="#M.Device"/>
  <owl:disjointWith rdf:resource="#Equipment"/>
  <owl:equivalentClass rdf:resource="&ucsl;Sensor"/>
  <owl:equivalentClass rdf:resource="&pal;Sensor"/>
</owl:Class>
```

```
<owl:Class rdf:ID="Bomb">
  <rdfs:subClassOf rdf:resource="#Weapon"/>
  <rdfs:subClassOf rdf:resource="#ExplosiveDevice"/>
  <owl:equivalentClass rdf:resource="&sumo;Bomb"/>
  <owl:equivalentClass rdf:resource="&pal;Bomb"/>
</owl:Class>
```

```
<owl:Class rdf:about="&ucsl;Equipment">
  <rdfs:subClassOf rdf:resource="#Device"/>
</owl:Class>
```

```
<owl:Class rdf:about="&meped;Equipment">
  <rdfs:subClassOf rdf:resource="#Device"/>
</owl:Class>
```

Example 3 Problem

- Models differ in what they define as first-class objects, and what is simply represented by name/code strings
 - Model A says *Location* is a first-class object, linked to by other objects
 - Model B says *Location* is represented in other objects by its name or country code string
 - How to relate the two?

Example 3 Solution: OWL 2's PropertyChainAxiom

- a's *countryOfBirth* is an ObjectProperty linking *Person* to *Country*
- b's *birthCountry* is a DatatypeProperty of *Person* simply naming a country

```
<rdf:Description rdf:about="&b;birthCountry">  
  <owl:propertyChainAxiom rdf:parseType="Collection">  
    <owl:ObjectProperty rdf:about="&a;countryOfBirth"/>  
    <owl:DatatypeProperty rdf:about="&a;name"/>  
  </owl:propertyChainAxiom>  
</rdf:Description>
```

Example 4 Problem

- Capturing security, provenance, confidence, etc is critical
- Agreeing on how to represent that is needed for secure handling and effective data sharing
- ICS500-21 requires all IC XML representations to use the ISM XML schema for security markings
- This is not workable for many pre-defined XML schemas such as RDF/XML
- The IC needs to agree on an ontology representation which can be ratified as a standard in lieu of ISM XML

Example 4 Solution: OWL 2's AnnotationProperty

```
<owl:Class rdf:ID="CVE_Classif">
  <owl:oneOf rdf:parseType="Collection">
    <owl:Thing rdf:about="#U">
      <rdfs:comment>UNCLASSIFIED</rdfs:comment>
      <ism:security rdf:resource="#U-USA"/>
    </owl:Thing>
    ...
  </owl:oneOf>
</owl:Class>

<ism:Security rdf:ID="U-USA">
  <ism:classification rdf:resource="#U"/>
  <ism:ownerProducer rdf:resource="#USA"/>
</ism:Security>

<owl:Class rdf:ID="Security">
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Restriction>
      <owl:onProperty rdf:resource="#classification"/>
      <owl:cardinality>1</owl:cardinality>
    </owl:Restriction>
```

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#ownerProducer"/>
  <owl:minCardinality>1</owl:minCardinality>
</owl:Restriction>
<owl:Restriction>
  <owl:onProperty rdf:resource="#classification"/>
  <owl:allValuesFrom rdf:resource="#CVE_Classif"/>
</owl:Restriction>
<owl:Restriction>
  <owl:onProperty rdf:resource="#ownerProducer"/>
  <owl:allValuesFrom rdf:resource="#CVE_OP"/>
</owl:Restriction>
<owl:Restriction>
  <owl:onProperty rdf:resource="#relTo"/>
  <owl:allValuesFrom rdf:resource="#CVE_Rel"/>
</owl:Restriction>
...
</owl:intersectionOf>
</owl:Class>
```

Example 4 Solution: OWL 2's AnnotationProperty

```
<owl:Axiom>  
  <owl:annotatedSource rdf:resource="#ID1"/>  
  <owl:annotatedProperty rdf:resource="#pdes;memberOf"/>  
  <owl:annotatedTarget rdf:resource="#ID2"/>  
  <ism:security rdf:resource="#Sec1"/>  
</owl:Axiom>
```

```
<ism:Security rdf:ID="Sec1">  
  <ism:classification rdf:resource="#S"/>  
  <ism:ownerProducer rdf:resource="#USA"/>  
  <ism:disseminationControls rdf:resource="#REL"/>  
  <ism:relTo rdf:resource="#USA"/>  
  <ism:relTo rdf:resource="#CAN"/>  
  <ism:relTo rdf:resource="#GBR"/>  
  <ism:relTo rdf:resource="#ISAF"/>  
</ism:Security>
```

- Asserts the triple
- Annotates the triple with its security

Example 5 Problem: Enumerations with Patterns

```
<cve:Enumeration multivalue="true" ism:ownerProducer="USA" ism:classification="U">
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA">RD</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA">RD-CNWDI</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA" regularExpression="true">RD-SG-[1-9][0-9]?</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA">FRD</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA" regularExpression="true">FRD-SG-[1-9][0-9]?</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA">DCNI</Value>
  </Term>
  <Term ism:classification="U" ism:ownerProducer="USA" >
    <Value ism:classification="U" ism:ownerProducer="USA">UCNI</Value>
  </Term>
</cve:Enumeration>
```


Example 5 Solution: OWL 2's use of rdfs:Datatype

```
<owl:unionOf rdf:parseType="Collection">
  <rdfs:Datatype>
    <owl:oneOf>
      <rdf:List>
        <rdf:first rdf:datatype="&xsd:string">RD</rdf:first>
        <rdf:rest>
          <rdf:List>
            <rdf:first rdf:datatype="&xsd:string">RD-CNWDI</rdf:first>
            <rdf:rest> ... </rdf:rest>
          </rdf:List>
        </rdf:rest>
      </rdf:List>
    </owl:oneOf>
  </rdfs:Datatype>
  <rdfs:Datatype>
    <owl:onDatatype rdf:resource="&xsd:string"/>
    <owl:withRestrictions rdf:parseType="Collection">
      <rdf:Description>
        <xsd:pattern rdf:datatype="&xsd:string">RD-SG-[1-9][0-9]?</xsd:pattern>
      </rdf:Description>
    </owl:withRestrictions>
  </rdfs:Datatype>
  <rdfs:Datatype> ... </rdfs:Datatype>
</owl:unionOf>
```

Future Work

- Add mappings in METS ontology to the concepts in UCore SL
- Continue retrofitting OWL 2 constructs into the ontologies
- Work with the Intelligence Community and the DoD on standardizing ontologies such as ISM

Contact Info

Richard Lee lee_richard@bah.com 571-482-7809

METS info and ontologies <http://purl.org/mets>