Semantic Information Modeling for Federation

OMG Architecture Ecosystem / ADTF

April 2016

Cory Casanave  (cory-c at modeldriven dot com)
What is SIMF?

• Semantic Information Modeling for Federation is an in-progress standards process within OMG

• The goal of SIMF is to provide the modeling capabilities to support information federation by leveraging conceptual and logical information modeling with model bridging relations

• An important goal of SIMF is to extend the economic lifecycle of many existing running systems that were conceived and built in a time when there were insufficient true conceptual capabilities, by explicitly adding the often implicit semantics and linking the running data representations with a true conceptual model.

• Another important goal of SIMF is not to replace any existing standard language but to build bridges between the SIMF Conceptual Schema and the various standard languages such that the users can save substantial amounts of money in their integration and federation efforts.

• Issued Dec 2011 - OMG Document: ad/2011-12-10
Proposition

Problem statement

- **Federation** (information sharing, interoperability, shared services, etc.) is the problem of this decade – it is costing productivity, lives and billions trillions of dollars annually. It is the pre-requisite to solving many problems in the large. It is a problem faced by most CIOs in government and industry.

- We are calling this the “data problem”

A problem not solved...

- **None of the standards we have directly target this problem.** Not: UML, OWL, LoD, E/R, SOA, DoDAF, XML Schema, Common Logic or SBVR, etc.

- With all these solutions – we still have a pervasive problem!

- **While not ideal, the standards above can and are used for federation, but, they are all built for other purposes and repurposed to solve the data problem.** Experts can pull these technologies together to solve a specific problem, we want to make it easy to do so with an integrated and standardized approach supporting mainstream solutions and internet-scale federation.

We can make a substantial dent in the data problem with new standards derived from current technologies and practices. This is the “SIMF” Initiative.
SIMF Principles

• The intent of SIMF is not to replace other standards but to make it economically more feasible that various standards can cooperate.
• In that sense SIMF is not a competitor to any existing OMG standard.
• It is the aim of SIMF to extend the economic life of existing applications as well as to support new applications.
• It is the intention of the submission team to extensively test the proposed standard set of concepts and languages before submission and make the test results available.
• SIMF does use conceptual models (reference models) as “pivot points” between different representations.
• SIMF does assume that some concepts will be agreed on by some communities as some agreement is required for any communication.
• SIMF does not assume any universal or all encompassing conceptual model.
Pivoting through a conceptual model

The conceptual pivoting approach

- A common and growing approach to the data problem leverages abstraction: Defining a domain focused vocabulary with integrity rules and assertions as part of a conceptual model that captures domain semantics. Federation and integration is achieved by relating various logical and physical information structures to the conceptual model.

- Information federation and integration is achieved via a “pivot” through this conceptual semantic layer.

- This approach is used, in part, in existing standards such as CCTS (Core Components), ISO 20022 and is currently being utilized in OMG for finance.

- In the majority of cases the “tool” used to represent these common semantics and links is a spreadsheet, but UML and OWL are also used.
Example of “Pivoting” through a conceptual model

There is an actual “Person”, Cory Casanave

- There is a concept of this person shared in this room, right now
- Here is one representation of him
- “Person” is a shared concept, independent of data structures
- There may also be shared agreement that Cory is a person and some other “facts”
  - “Cory Casanave” is a name for this person
  - He weighs 240 LBS
- There are multiple data representations about Cory Casanave which may or may not agree
- Those representations can be grounded in concepts (semantics), assisting federation
Subject focused conceptual models define the concepts, predicates, integrity rules and terms of a domain that can be related to each other.

Solution focused logical information elements represent information structures and integrity rules that can use and extend other information.

Technology focused physical data schema are grounded in logical data models which define their context and semantics.
SIMF Components – conceptual models and mappings

- Conceptual Modeling MetaModel
- Mapping MetaModel
- Schema Model

- Conceptual Modeling UML Profile
- Mapping UML Profile
- Schema Model (From IMM)

- UML Representation

- Depends On
- Map To
## Comparing SIMF and OWL

<table>
<thead>
<tr>
<th><strong>SIMF</strong></th>
<th><strong>OWL</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual models defined for federation</td>
<td>Ontologies defined for general inference</td>
</tr>
<tr>
<td>Focus on understandable domain concepts</td>
<td>Focus on inference</td>
</tr>
<tr>
<td>Includes mapping language - Assumes data is &quot;messy&quot;</td>
<td>Has no mapping language - Assumes data is consistent</td>
</tr>
<tr>
<td>N-Ary first class relationships</td>
<td>Directed “triples” without identity</td>
</tr>
<tr>
<td>Provides for common restrictions and assertions</td>
<td>Provides more general restrictions and assertions</td>
</tr>
<tr>
<td>Context as a first-class concept</td>
<td>Limited to first order concepts</td>
</tr>
<tr>
<td>Intended to cross viewpoints</td>
<td>Intended to represent a viewpoint</td>
</tr>
<tr>
<td>UML Profile understandable</td>
<td>UML Profile difficult</td>
</tr>
<tr>
<td>Defines needed domain “meta concepts” like roles, phases and units</td>
<td>Common domain meta concepts are not part of the language</td>
</tr>
<tr>
<td>OWL can be generated and augmented</td>
<td>OWL is base language</td>
</tr>
</tbody>
</table>
Example Snippet

**Custody**

The act of protecting or taking care of something [merriam-webster.com]

**Custodian**

An actor who has responsibility for or looks after some managed entity. A Custodian has custody of a managed entity via the Custody relation.

**Managed Entity**

Any entity for which the custody of or access to the entity is managed such that it can be trusted or protected. A managed entity is in the custody of a custodian via the Custody relation.