Discovery and Reuse of Modeling and Simulation Assets / Paper 10S-SIW-048

Gustavson, Paul

http://hdl.handle.net/10945/30782
ABSTRACT: The ability to discover existing modeling and simulation (M&S) assets is a critical need for enabling effective reuse and for reducing the duplication of capabilities. Such visibility and accessibility is key to optimizing the investment of the estimated billions of dollars spent on M&S within the Department of Defense (DoD).

This paper updates the status of work presented at the SISO 2009 Spring SIW (09S-SIW-076). The DoD M&S Community of Interest (COI) Discovery Metadata Specification (MSC-DMS) version 1.2.1, which provides a common mechanism to catalog a wide variety of M&S resources as metacards, is now publicly available through the DoD Metadata Registry. A formal process has been established to coordinate the MSC-DMS and the closely related specification used by the DoD M&S Catalog.

The M&S Catalog provides a common portal with connection to various M&S related repository sources, and is expected to be declared operational in early 2010. In addition, the M&S Catalog will be federated with the DoD Enterprise Catalog, and publish-and-subscribe mechanisms using DoD Net-Centric Enterprise Services (NCES) will be developed. Tools will be developed and technical assistance provided to help sources map their existing data structures to the MSC-DMS, to generate conformant XML-formatted metadata, and to submit the metadata to the M&S Catalog (or any other authorized subscriber). Potential sources of metadata are encouraged to contact the authors.

Disclaimer: The views presented in this paper are those of the authors and do not necessarily represent the views of the Department of Defense or its Components.

1. Introduction

This is the third installment of a technical paper within SISO regarding the M&S Community of Interest (COI) Discovery Metadata Specification (MSC-DMS). The first paper, presented in the Fall of 2008, was simply an introduction; a primer describing how it was developed, what it offered including what types of resources could be cataloged and what available metatags could be used to do it [1].

The second paper, presented in the Spring of 2009, focused on maximizing discovery; it introduced the M&S Catalog, which provides a common portal with connection to various M&S related repository sources [2]. This second paper highlighted what could be exploited from the MSC-DMS by search engines such as the one provided by the M&S Catalog.

This third paper focuses on the practical application of the specification and use of the M&S Catalog. Like the second paper, the MSC-DMS and M&S Catalog are featured as a dynamic duo. However we further highlight what the MSC-DMS and M&S Catalog both presently offer to the community, and examine how they can be used in a practical way for both developers in creating metacards, that describe their resource
(similar to a soup can label describing a food product), and users who are looking to discover such resources.

1.1 Net-Centric M&S Discovery Efforts

The organization that has supported the development of the MSC-DMS is the Modeling and Simulation Coordination Office (M&S CO), which is chartered by the Office for the Secretary of Defense (OSD). The first version of the MSC-DMS appeared on the street almost two years ago. In those two years it has been updated twice and uploaded to the DoD Metadata Registry [3]. In addition, a supporting document, identified as the MSC-DMS Implementation Guide [4] has recently been developed that describes how to document and catalog M&S resources with the necessary Discovery Metadata. The goal of both the MCS-DMS specification [5] and guide is to help support the net-centric goals of data visibility of M&S assets across the DoD.

Also in the past two years a growing number of people have begun to utilize the MSC-DMS as a template for creating metacards to catalog resources. A majority of these users are involved in connecting their resource repositories with the M&S Catalog effort, which was birthed just over a year ago. Users of the M&S Catalog and MSC-DMS include the Army, Navy, Air Force, and DoD MSRRs, and joint analysis community among others.

Twice now we’ve used the term metacard. Perhaps you are wondering what a metacard is. A metacard holds key information typically in Extensible Markup Language (XML) format that describes a resource including its purpose and application, and other information including points of contact, creation date, and, if available, usage experience. We say XML, because that’s how it is supported by the MSC-DMS, and also by its close relative, the DoD Discovery Metadata Specification (DDMS) [6], which provides the basis for all community focused discovery metadata initiatives within DoD.

In support of the different versions of the MSC-DMS that have evolved and to maintain mapping with the DDMS, several transformation files based on the Extensible Stylesheet Language Transformation (XSLT) have been developed as well. These XSLT files help transform MSC-DMS metacards to support different specified formats. Specifically they help maintain compatibility with different versions of the MSC-DMS as they have been introduced, and also cross compatibility with the DDMS thereby supporting the DoD Core Enterprise Services such as DoD Federated Search and Enterprise Catalog (see side note).

The bottom line is that in the last two years the tools have been put in play to begin to better build, connect, and share M&S resources. These tools include the MSC-DMS, supporting guides, examples, transformations, the M&S Catalog, and the connection of repositories to the M&S Catalog. With all these things in place organizations are poised to help improve M&S reuse and discovery – but only if the capabilities are used and used properly. Furthermore, these capabilities stand to better serve the needs of the DoD M&S community only if feedback continues to be received that identifies how to improve these tools.

Note: The purpose of DoD Federated Search is to provide a capability to the community of users (including but not limited to M&S users), which allows the discovery of information from disparate data sources with “one query” request. Such a Federated Search request works with the Enterprise Catalog, which uses standard, vendor-neutral specifications to leverage existing Community of Interest (COI) information repositories such as the M&S Catalog.

The Enterprise Catalog makes connection with a network of resources to support the Federated Search goal. Understanding that each COI has different needs with regard to how and what should be reflected in its metadata (e.g., the MSC-DMS for DoD M&S resources) but knowing that a core lingua franca (e.g., DDMS) is needed for the federated search clarifies why transformation mappings, such as described above, are needed. For more information on these Core Enterprise Services visit https://metadata.dod/mil/mdr.

1.2 M&S Assets to Discover

Discovery is defined as “the ability to locate data assets through a consistent and flexible search.”[6] The DoD Net-Centric Data Strategy (dated May 9, 2003) [7] defines goals and approaches for users and systems to
discover and access a wide range of data assets throughout the DoD Enterprise.

From a MSC-DMS position, the types of M&S resource that should be cataloged with the same type of metacards include the following:

1. M&S software (implements a model or simulation)
2. Adjunct tools (e.g., data loggers)
3. Federations of simulations
4. M&S software components
5. M&S services (models and simulations implemented as web services)
6. M&S data (data in M&S-usable format and data produced by M&S)
7. M&S data models (structural metadata for M&S data)
8. Interface specifications
9. M&S software design documents

// probably could roll into how the Catalog supports the discovery of these metacards.

// describe how the M&S Catalog works

Represented in Figure Y is a diagram of the processing of metadata received for the M&S Catalog.

// do we need to reference it as part of the M&S Catalog how-to guide for sources

Each Source identified in this illustration represents a resource repository. They are sometimes referred to as a data source. It shares with the M&S Catalog a record of the resources it maintains. The M&S Catalog takes that information, validates it, and then provides a clearinghouse listing of what resources are available from these other sources to users. The M&S Catalog maintains discovery metadata, and the data that the M&S Catalog provides to users are the metadata elements defined in the MSC-DMS [5].

If possible, as much as possible of the metadata content in the M&S files managed by the data source that corresponds to the MSC-DMS data elements are to be included in the data provided by a Source to the M&S Catalog. At a minimum the metacard data provided by the Sources should include:

- Resource ID
- Title
- Description
- POC

Certain other fields are highly recommended but not absolutely required for Source submission by the M&S Catalog. To be compliant with the MSC-DMS schema, the following fields should also be present:

// could talk briefly about how M&S Catalog supports reflecting these metacards, and also other asset types including SMEs.

The overall category of elements that the MSC-DMS provides for documenting a metacard is illustrated in Figure X.

// Standard Figure goes here showing the basic template of the MSC-DMS

The M&S Implementation Guide explains how to document and capture these metacard elements. It also provides a basic XML tutorial to help the reader understand the syntax. Some of the XML syntax is used within this paper, and the reader is encouraged to leverage the Implementation Guide or other XML related references if such syntax is not clear.

The M&S Catalog team has identified that there are subcomponents of the MSC-DMS that also need to be cataloged and discovered. These include subject matter experts (SMEs), which can be captured using the POC type provided by the MSC-DMS. POCs include individuals (identified as Person) and Organizations. The overall category of POC elements for these two types is illustrated in Figure Y.

// POC type template goes here
Very important is the fact that the M&S Catalog explicitly wants to stay as closely aligned with the MSC-DMS as possible. It hasn't been perfect because the M&S Catalog has been on the "cutting edge" supporting Sources with legacy metadata and assimilating into something that fits as closely as it can with the MSC-DMS. But the vision and the process are to keep them going in the same direction.

2. Generating and using IDs within MSC-DMS

There are several components of the MSC-DMS that utilize unique identifier (ID) tags to support cross referencing. They include the following:

- Resource
- Associations
- POCs (Organizations and Persons)

IDs can be very useful for uniquely marking assets and referencing other previously marked assets. The above three uses of IDs are discussed in the following subparagraphs.

2.1 Resource

A complete resource record may be marked by a unique identifier to support cross referencing by other resources and for the benefit of organizing data by one or more repositories. The ID provides a way to uniquely distinguish a resource allowing it to be referenced and used by community members. In fact, with version 1.2 – there can be two types of IDs identified: an ID for the metacard, which stands alone from the resource; and an ID for the resource itself. In theory the ID values could be the same; however, the M&S Catalog has elected to differentiate between the metacard and the resource by using a different ID. Fortunately the metacard can be used to identify the related resource ID.

An example of an ID is shown in the following listing.

In this XML snippet, the ms:Resource tag identifies a unique ID in the ms:ID attribute, which relates to the specific M&S resource. The ms:Metacard_ID attribute provides a unique ID that can be associated to the standalone Metacard represented by this XML metacard document. The attributes that follow identify the related name spaces used to validate the content of the XML metacard document.

2.2 Associations

An MSC-DMS metacard can define associations to other M&S resource assets or support assets. This is a very powerful component for supporting Semantic Publishing, which is discussed in greater detail later. However, an XML metacard example is provided below that shows how associated resources are referenced by ID.

In this XML snippet, the ms:Resource tag identifies an Association to another document. An ID Tag (ms:associationID attribute) of “4352” is used to mark what additional resource it connects to. We will explore this further later in this paper.

2.4 Points of Contact (POCs)

IDs are also used to identify POC organizations and persons. The types of IDs for POC organization and persons are shown in Table X.
Both an Organization and Person are uniquely identified by an ID. It is intended to separate that organization or person from others.

### 2.4.1 POC.Organization

An organization may identify what sponsor it is supporting pertaining to the resource of interest. This is marked with the sponsorID attribute. Additionally the organization may also identify a parent organization that it is a part of. An example highlighting how to document a POC Organization element is provided below:

```xml
<ms:POC>
  <ms:Role ms:value="publisher"/>
  <ms:Organization ms:ID="330" ms:parentID="320" ms:sponsorID="720">
    <ms:Name ms:value="SpyFlight IT"/>
    <ms:Type ms:value="industry"/>
    <ms:AddressInfo>
      <ms:Phone ms:type="work" ms:number="540-324-2208"/>
      <ms:Email ms:type="work" ms:address="info@spyflt.com"/>
      <ms:URL ms:value="http://www.spyflt.com"/>
      <ms:ContactInstruction ms:value="For general assistance, dial 0 for an operator"/>
    </ms:AddressInfo>
  </ms:Organization>
</ms:POC>
```

In this XML snippet, the metacard identifies a POC Organization uniquely identified by the value “330” in the ms:ID attribute. Assuming this value is somehow registered with that organization (via a repository for example), then the ID value can be used as an association or reference by other POCs. For example, this XML snippet references a sponsor with the ID tag “720” (ms:sponsorID attribute), and references a parent organization with the ID tag “320” (ms:parentID attribute).

### 2.4.2 POC.Person

Like an Organization, a Person may identify what sponsor he or she is supporting pertaining to the resource of interest. This is marked with the sponsorID attribute. A Person may also identify his/her supervisor (i.e., manager) pertaining to the resource of interest. An example highlighting how to document a POC Person element is provided below:

```xml
<ms:POC>
  <ms:Role ms:value="publisher"/>
  <ms:Person ms:ID="53530" ms:sponsorID="720" ms:supervisorID="51230">
    <ms:Title ms:value="Technical Reviewer"/>
    <ms:Name ms:first="Lindsey" ms:last="Piddleton"/>
    <ms:Position ms:value="Writer">
      <ms:Org ms:organizationID="320">
        <ms:Name ms:value="SpyFlight Industries"/>
      </ms:Org>
    </ms:Position>
    <ms:AddressInfo>
      <ms:Phone ms:type="mobile" ms:number="540-755-5555"/>
      <ms:Email ms:type="work" ms:address="writer@spyflight.com"/>
      <ms:ContactInstruction ms:value="Leave message w/ personal assistant if not available."/>
    </ms:AddressInfo>
  </ms:Person>
</ms:POC>
```

In this XML snippet, the metacard identifies a POC Person, Lindsey Piddleton, uniquely identified by the value “53530” (ms:ID attribute). Assuming this value is somehow registered to that person (i.e., last 5 digits of her social security card for example), then the ID value can be used as an association or reference by other POCs. For example, this XML snippet references a sponsor with the ID tag “720” (ms:sponsorID attribute), and references a supervisor (i.e., manager, or TPOC) with the ID tag “51230” (ms:supervisorID attribute).

### 2.4 Generating and Using Unique IDs

One of the tools available to help define a unique ID is a web service known as UUID Generator (http://www.uuidgenerator.com/). Neither the MSC-DMS, nor the Implementation Guide nor M&S CO, which has supported the MSC-DMS development effort, mandates the use of this service or any other...
similar service. But it is something to consider. Keep in mind though it should be the responsibility of the individual and his/her organization to identify by what means an ID should be distributed and received. It is likely that the policy and practice for attaining IDs would be provided or addressed by the repository and repository portals for which an individual and organization wishes to participate.

For purposes of the M&S Catalog, the Resource ID character string should be unique and stable for the M&S Catalog. It should be unique in the sense that at any given time, no two records from a single Source, which maintains M&S resources such as repository, should have the same Resource ID value. And it should be stable in the sense that if a metacard record is submitted that updates or replaces a metacard record submitted earlier, then it should have the same Resource ID value as the earlier metacard record.

// Richard – do you want to add anything else here regarding the M&S Catalog’s policy and practice for IDs here.

3. Creating Semantic Meaning via Associations

No matter what type of M&S resource may be cataloged, invariably there are many other materials that could be associated with the resource. The MSC-DMS provides a powerful mechanism to capture and reflect these associations. The MSC-DMS and Implementation Guide walks through the each of the specifics that can be attributed to an Association, but we highlight some of the useful characteristics to support the concept of Semantic Publishing.

First we should explain what we mean by Semantic Publishing. The term comes from the Semantic Web community. Semantic Publishing refers to publishing information on the web as document resources accompanied with semantic markup. Publishing resources in such a way allow computers to digest the markup and understand the meaning of published information thereby facilitating better search and data integration.

The Associations component of the MSC-DMS provides a way to correlate related resources and artifacts. Specifically it supports the following type of relationships:

- has-a
- is-part-of
- is-type-of
- is-described-by

It’s important to understand what these four relationship types mean.

- Has-a indicates that the M&S resource is composed of another asset. This type of association therefore identifies the child to that resource. It is useful for supporting decompositions.
- Is-part-of indicates that the M&S resource is part of the composition of another asset. This type of association therefore identifies a parent asset. It is useful for supporting compositions of aggregate sets.
- Is-type-of indicates that the M&S resource is an instance or a manifestation of another type of asset. This type of association therefore identifies the type of asset the resource exemplifies. It is useful for resources that support a specific need, specification or use case.
- Is-described-by indicates that the M&S resource is specified or more clearly described by another asset. This type of association identifies the clarifying document, specification, material (such as a web site), or subject matter expert (SME). It is useful for providing more amplifying information by providing a link to such information rather than having to restate specific information within the metacard.

Type of associations that be identified within version 1.3 of the MSC-DMS includes the following:

- software
- tools
- federations
- software components
- services
- data
- data models
- interface specifications
- software design documents
- infrastructure
- supported events
- future capabilities
- requirement
- related documents
- environment
- subject matter experts (new)
The nine items on the left hand column are reflective of the types of things that not only can be associated to an MSC-DMS metacard, but the types of artifacts that could be fully cataloged and described by the MSC-DMS template itself. The items on the right hand column reflect the types of things that are not intended to be cataloged or described by an MSC-DMS metacard, but which still may be referenced by a MSC-DMS based resource. However, it should be noted that a Subject Matter Expert, which has been added as new enumeration value for the pick list of association types in Version 1.3, can be cataloged and described by a subcomponent of the MSC-DMS known as the POC type. POC type includes both Organization and Persons.

Multiple associations can be made by the MSC-DMS resource metacard. This is illustrated in the XML snippet provided below for a metacard describing a notional simulation, which models surface-to-surface missiles:

```xml
<ms:Associations>
  <ms:Association
ddms:qualifier="MIL-STD-498"
ddms:value="Ballistic Flight Model SRS"
ms:relationship="is-type-of"
ms:type="requirement"
ms:associationID="8A3EFF336"
ms:constraints="exclude section 4.6"/>
  <ms:Association
ddms:qualifier="C++"
ddms:value="Ballistic Scud Model Source Code – Al Hussein"
ms:relationship="has-a"
ms:type="software_components"
ms:associationID="73343323"/><ms:Association
ddms:qualifier="C++"
ddms:value="Ballistic Scud Launcher Module"
ms:relationship="has-a"
ms:type="software_components"
ms:associationID="247098233d"/>
  <ms:Association
ddms:qualifier="HLA"
ddms:value="Missile Defense HLA Federation"
ms:relationship="is-part-of"
ms:type="federations"
ms:associationID="f0923f93"/>
  <ms:Association
ddms:qualifier="URL"
ddms:schemaHref="http://www.shortrangeschemas.com"
ms:relationship="is-described-by"
ms:type="related document"
ms:associationID="4352"/>
</ms:Associations>
```

In this notional XML metacard excerpt, five associations are shown. The first association identifies a requirement or set of requirements as depicted by a MIL-STD-498 document identified as the “Ballistic Flight Model SRS” with the M&S resource; indicating that the resource is-type-of Ballistic Flight Model Simulation. The second and third association, which are has-a relationship types, marries software components with the resource metacard. The fourth association identifies that the simulation is-part-of an HLA federation. And finally, the fifth association identifies a related HTML document that describes more information.

The specification and implementation guide provide further amplification of the Association element regarding how other attributes and sub elements, which includes a Description component, can be used.

// could talk about taxonomies here and creating / establishing better – more well understood meaning (to support Semantic Web), which is discussed in Paul G’s other SIW Paper. This other SIW Paper identifies a road map toward Semantic Web – that could start with exploiting the MSC-DMS – specifically Associations.

4. What’s New in MSC-DMS 1.3

// Short discussion here at what’s be most recently added to MSC-DMS.

Since the last time the MSC-DMS was highlighted to the community, the MSC-DMS has benefitted from 2 minor yet significant updates: Version 1.2.1 and Version 1.3, which is anticipated to be released by Spring 2010 SIW. Changes have included update to several types of enumerated values (or “pick-list” as it is commonly referred) that are available. The specific pick-lists that have been updated that can be assigned to some of the MSC-DMS attribute values include the following:
• Usage Application Domain Set (added “intelligence”)
• Date Type (Added “last_verified”)
• Association Qualifier (was open text field, added new pick-list which includes “URL”)
• Association Type (added new association types including “Subject Matter Expert”)

In addition to these enumeration types, the following other adjustments have been made for version 1.3:
• Formerly added a capability to identify specific taxonomies for any MSC-DMS element
• Modified and improved the VV&A Coverage Metadata Set extension/supplemental component
• Corrected grammatical errors (such as PostalCode, which was misspelled as PostcalCode)

// this probably needs to be iterated some more.

5. Summary

The ability to discover existing modeling and simulation (M&S) assets is a critical need for enabling effective reuse and for reducing the duplication of capabilities. The MSC-DMS and M&S Catalog provide two tangible and related tools for facilitating discovery and reuse. The DoD community is encouraged to leverage both these tools.

// recap more here

References
[1] Fall 2008 paper
[2] Spring 2009 paper
[3] DoD MDR
[7] Net-Centric Data Strategy

Author Biographies

PAUL GUSTAVSON is Chief Scientist at SimVentions, Inc, and has over 19 years of experience including the design and development of DoD standards, simulation systems, software applications, and numerous technical publications and tutorials on simulations and software development. Currently he is supporting the M&S Coordination Office (M&S CO) in helping to identify key metadata needed for the cataloging and discovery of M&S assets. He has been an active leader within the Simulation Interoperability Standards Organization (SISO) involved in multiple standards efforts including the Base Object Model (BOM), Distributed Simulation Engineering and Execution Process (DSEEP), and HLA Evolved. He is a co-author of several books including “C++Builder 6 Developer’s Guide.” Mr. Gustavson holds a Bachelor of Science degree in Computer Engineering from Old Dominion University, and lives in Virginia with his wife and two boys.

CURTIS BLAIS is a research associate in the Naval Postgraduate School Modeling, Virtual Environments, and Simulation (MOVES) Institute. He has over 35 years experience in M&S research, development, and education. Mr. Blais provides technical consultation in web-based technologies for development and employment of the M&S Catalog and is coordinating the generation and transfer of metadata describing 3D models maintained in the MOVES Scenario Authoring and Visualization for Advanced Graphical Environments (SAVAGE) repository. He is an active member of the SISO Product Development Groups for the Coalition Battle Management Language, Military Scenario Definition Language, and Simulation Reference Markup Language. Mr. Blais holds Bachelor of Science and Master of Science degrees in Mathematics from the University of Notre Dame.