M&S Asset Discovery: Services, Tools and Metadata

Paul Gustavson
SimVentions, Inc.
Fredericksburg, VA
pgustavson@simventions.com

Richard Daehler-Wilking, Ph.D.
Space and Naval Warfare Systems Center
Charleston, South Carolina
r.daehler-wilking@navy.mil

Curtis Blais
MOVES Institute, Naval Postgraduate School
Monterey, CA
clblais@nps.edu

Hart Rutherford
SimVentions, Inc.
Fredericksburg, VA
hrutherford@simventions.com

ABSTRACT

Modeling and Simulation (M&S) has become an extremely effective tool for supporting modern warfighting. This includes the use of M&S in systems acquisition and test and evaluation to improve the engineering performance of equipment and weapons systems, the use of M&S in training warfighters in essential decision-making and operational skills, and the use of M&S for analysis and experimentation to determine and evaluate future force capabilities. For each of these applications and others, cost-effective development and employment of M&S is crucial to U.S. warfighting success.

One way to achieve cost-effectiveness is to maximize use of existing M&S assets when those assets can meet requirements, rather than by initiating new developments that run the risk of failing to achieve such requirements. Maximizing use of existing assets, the essence of re-use, necessitates the development of mechanisms that enable systems acquisition personnel, systems development personnel, and operational users to discover and access the appropriate assets for their particular needs. Discovery of existing M&S assets is essential to reuse, but requires human and computer-understandable descriptions of their scope, features, and application to be effective. While thousands of entries exist in various M&S repositories and registries, maintaining their currency and completeness is a long-standing problem. To address the problem, the DoD Modeling & Simulation Coordination Office (M&S CO) is providing discovery services, tools, metadata specifications, and policy and guidance. This paper reviews these activities and how they can help generate, communicate, and make visible discovery metadata about M&S assets of interest to the DoD.

These activities center upon the M&S Catalog (https://mscatalog.osd.mil) that connects asset owners, repositories, and disparate end-users. Cloud-based tools such as the Enterprise Metacard Builder Resource (EMBR) Portal support collaborative creation of resource metacards. Standards such as the DoD M&S Community of Interest (COI) Discovery Metadata Specification (MSC-DMS) provide the template for documenting M&S assets. The newly revised DoD M&S Glossary (DoD 5000.59M) provides a clear M&S vocabulary and taxonomy for community understanding. Newly defined policy and guidance (such as DoDI 5000.59) help facilitate the sustainability of discovery and reuse of assets for the broader M&S community.

Version 1.4 of the MSC-DMS includes enhanced support of taxonomies, and allows a metacard to say exactly where individual terms come from (e.g., the M&S Glossary). Additionally, the MSC-DMS now supports records of points of contact as stand-alone entries. To assist potential contributors, the M&S CO has stood up the Discovery Data Working Group (DDWG) that provides a forum for mutual support and ongoing improvement of all aspects of discovery.

Together, these services, tools, standards, and policies and working groups promise to help further the discovery, reuse, and interoperability of M&S well into the 21st century.

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.

ABOUT THE AUTHORS

PAUL GUSTAVSON is Chief Technology Officer and co-founder of SimVentions, Inc. He has over 22 years of experience including the design, development, and integration of DoD systems, simulations, standards, software applications, and has authored numerous technical publications and tutorials on simulations and software development. Currently he is supporting 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. He is also a certified John Maxwell Team coach and speaker.

RICHARD DAEHLER-WILKING has been a computer scientist in the Command and Control Department of the Space and Naval Warfare Systems Center Charleston, South Carolina, since 2000. He was a longtime participant in the NETWARS/JCSS community, having led the development of models for both EHF Milstar and UHF DAMA satellite communications. More recently, he has supported the Navy M&S Office (NMSO) and now serves as technical lead for the DoD M&S Catalog project. He has a Bachelor of Arts degree in Mathematics from Reed College, a Master of Arts degree in Mathematics from the University of Oregon, and a Ph.D. from the Medical University of South Carolina in biometry (systems science track). He taught mathematics, computer programming, and physics at the college level for 25 years. His hobbies include alpine skiing and barbershop harmony.

CURTIS BLAIS is a Research Associate Professor in the Naval Postgraduate School's Modeling, Virtual Environments, and Simulation (MOVES) Institute. He has over 35 years of 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.

HART RUTHERFORD is the manager of the M&S portfolio of products and services at SimVentions Inc. Mr. Rutherford has over 20 years of professional experience as a combat systems engineer and program manager including technical leadership of M&S VV&A for the Navy's DD(X) program as well as a contributor to the development and maintenance of DoD M&S standards. Mr. Rutherford's military background includes 8 years active and reserve service in the U.S. Navy as an AEGIS Operations Specialist. He holds a Master's degree in Systems Engineering from Old Dominion University and B.S. in Computer Information Systems from Chapman University.

M&S Asset Discovery: Services, Tools and Metadata

Paul Gustavson
SimVentions, Inc.
Fredericksburg, VA
pgustavson@simventions.com

Richard Daehler-Wilking, Ph.D.
Space and Naval Warfare Systems Center
Charleston, South Carolina
r.daehler-wilking@navy.mil

Curtis Blais
MOVES Institute, Naval Postgraduate School
Monterey, CA
clblais@nps.edu

Hart Rutherford
SimVentions, Inc.
Fredericksburg, VA
hrutherford@simventions.com

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.

INTRODUCTION

Information drives decision making. Consider the world we live in. Visual and verbal information affects our buying decisions. As consumers we select what we buy based on either what we already know about a product, or how that product is labeled and marketed. Publishers have recognized this for decades — who doesn't judge a book (or a magazine) based on its cover? Likewise, grocers know that what's identified on the box helps sell the product.

Consumers look for information that meets their interest. For example, information that might help you assess whether or not you place a box of granola bars in your shopping cart might tell you if the product is (a) rich in fiber, (b) is a good source of calcium, (c) is low in fat, (d) contains whole grains, or (e) is gluten free. Plus if it's "new", then that's valuable information as well. For information to be effective it has to be accessible, visible, and understandable – it has to be right there in front of you in the form of easily digestible, discoverable metadata.

Discoverable information drives decision making at all levels of society – including the modeling and simulation (M&S) community. Consider that M&S practitioners are more apt to use products and tools, data models and standards, if the metadata (how each is described) is easily found and described to meet their needs. Metadata is data about data; it's data about an asset. More specifically metadata describes "characteristics of information" to aid in the "identification, discovery, and assessment" of a good or service.

M&S has become an extremely effective tool for supporting modern warfighting. This includes the use of M&S in systems acquisition and test and evaluation to improve the engineering performance of equipment and weapons systems, the use of M&S in training warfighters in essential decision-making and operational skills, and the use of M&S for analysis and experimentation to determine and evaluate future force capabilities. For each of these applications, and others, cost-effective development and employment of M&S is crucial to U.S. warfighting success.

One way to achieve cost-effectiveness is to maximize the use of existing M&S assets when and where it's applicable. For instance, when an M&S asset meets and supports requirements pertaining to a system acquisition, T&E, training, or analysis and experimentation need, then it should be considered for integration. The alternative is initiating new developments that run the risk of failing to achieve such requirements.

However, maximizing use of existing assets, the essence of reuse, necessitates the development of mechanisms that enable systems acquisition personnel, systems development personnel, and operational users to discover and access the appropriate assets for their particular needs. The ability to discover existing M&S assets is a critical need for enabling effective reuse and for reducing the duplication of capabilities. Metadata information describing an object of interest is the fuel that helps facilitate that discovery. However, historically, metadata for M&S assets has come in all shapes and formats or none-at-all. Inconsistency of the metadata, in what's provided and how it's presented, reduces the effectiveness of discovery. And yet, the visibility, accessibility, and understandability of

metadata are keys to optimizing the investment of the estimated billions of dollars spent on M&S within the Department of Defense (DoD).

Recent activities are making positive inroads within the M&S community and helping put the right information – the right labeling – out there. These include

- The advancement of a common, consistent metadata standard known as the M&S Community of Interest (COI) Discovery Metadata Specification (MSC-DMS);
- The establishment of a central access point, known as the M&S Catalog, which serves as a "shopping mall," providing access to various repositories (i.e., stores) with better labeled assets; and
- The emergence of a new set of tools that help producers and integrators better create and distribute labels for the M&S goods that they have developed, used, and should be made available to others.

Communities of interest that are affected by and/or that influence M&S include the data sources, data proponents, data curators and data repositories. These communities have the goal to document assets using tools and standards such as the MSC-DMS in order to share information about these assets through the M&S Catalog using common discovery metadata. That is, the M&S Catalog is built upon the needs and capabilities of the communities of interest through the application of metadata standards and tools, as reflected in Figure 1.

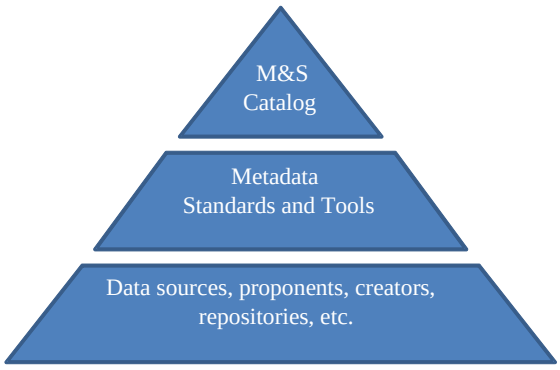

Figure 1: Data Discovery Pyramid of the M&S COI

If the discovery and reuse of M&S assets is important to you or your organization, then, whether your affiliation is government, industry, or academia, what truly matters to you is metadata. In the remaining text we will explore each of these aspects in greater detail.

METADATA STANDARDS (THE MSC-DMS)

When a query engine is searching thousands of documents, how does it know where to look for a title or description? What if there are multiple dates associated with a document? How does the query engine know when it has stumbled upon a creation date or a revision date?

In the future, when multiple query engines with different implementations are in use, how can useful metadata be produced without requiring the producer to generate a separate copy in a separate format for each such engine? One possibility would be for each search engine to convert each source's information into the format required by the search engine. However, this still requires creating a separate translation for every combination of source and destination, which is impractical in the long run. The key is to have consistent, well-marked metacards based on an agreed standard. The DoD MSC-DMS provides a Discovery Metadata specification suitable for use within the M&S COI. The components of the MSC-DMS enable a query engine, software program, or an individual to efficiently and effectively catalog and discover M&S assets of interest.

The MSC-DMS, which is available through the M&S Coordination Office (M&S CO) website [12] and the DoD Metadata Registry [6], is based on the DoD Discovery Metadata Specification (DDMS) [9] but tailored specifically for the M&S community.

The MSC-DMS supports the intent of the DoD Net-Centric Data Strategy as it relates to M&S. It is used to serve as a common mechanism to catalog a wide variety of M&S assets as metacards about M&S resources themselves, people and organizations (contacts) associated with those resources, and taxonomies that clarify terms used to describe the resources. The MSC-DMS allows metacards to be created for resources, contacts, and taxonomy classifications with the goal to help make them more visible, accessible, and understandable.

Resource Metacards

An M&S resource represents one or more assets that contribute to the composition or operation of an M&S event, environment, or infrastructure. A resource metacard holds key information that describes a resource including its purpose and application, and other information including points of contact, creation date, and, if available, usage experience.

The vocabulary and rules for identifying an M&S resource metacard are defined by the MSC-DMS within an Extensible Markup Language (XML) schema. The MSC-DMS XML schema ensures that an MSC-DMS resource metacard meets the rules defined in the specification. The following M&S resource types can be documented using an MSC-DMS resource metacard.

1. M&S software (implements a model or simulation)
2. Adjunct tools (e.g., data loggers)
3. Federations
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 model specifications
9. M&S support documents

Figure 2 provides an illustration of the MSC-DMS resource structure, which integrates core and supplemental layer component views pertaining to discovery metadata. The boxes with a solid outline represent required metadata components, whereas the boxes with a dashed outline represent optional metadata components.

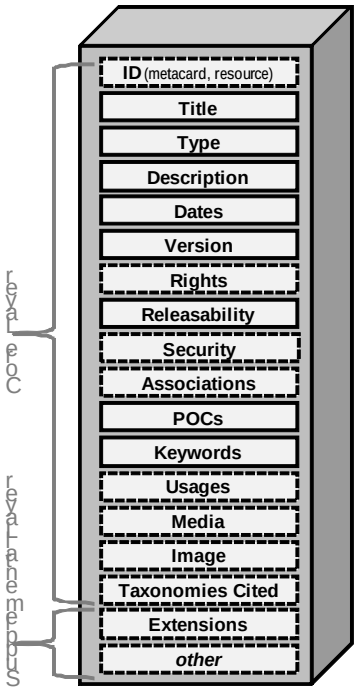

Figure 2: Resource Metacard Structure

The MSC-DMS document and the supporting Implementation Guide [11] provide greater details on how to begin to document a resource metacard. Additionally, the metacard tools, highlighted later in this paper, provide a useful mechanism to build these resource metacards.

Contact Metacards

Beginning with version 1.4, the MSC-DMS provides a means to capture and share contact information independently from a resource metacard. An M&S contact representing a person or an organization can help others achieve related goals such as reusing and/or integrating M&S resources. For instance, when simulation developers wish to pose detailed questions to those who were responsible for the development and distribution of a model, point-of-contact (POC) information can be vital. The POC information can be stored inside the record for the model itself or (as of v 1.4) as a stand-alone contact record. Another example might be a subject matter expert (SME) with experience using a variety of models and simulations in a variety of exercises. Rather than enter a copy of the SME's information into every single resource record (from a variety of sources), it can now be entered into a single contact record that lists his/her relevant expertise and experience. If "EADSIM" is listed in the description (or some other field) for the SME, then a text search for "EADSIM" should bring up both resource records about the EADSIM model itself and the contact record for the SME.

Figure 3 provides an illustration of the MSC-DMS contact structure, which is derived from the POC element of earlier versions of the MSC-DMS, and accommodates information about an individual person or organization. The boxes with a solid outline represent required metadata components, whereas the boxes with a dashed outline represent optional metadata components.

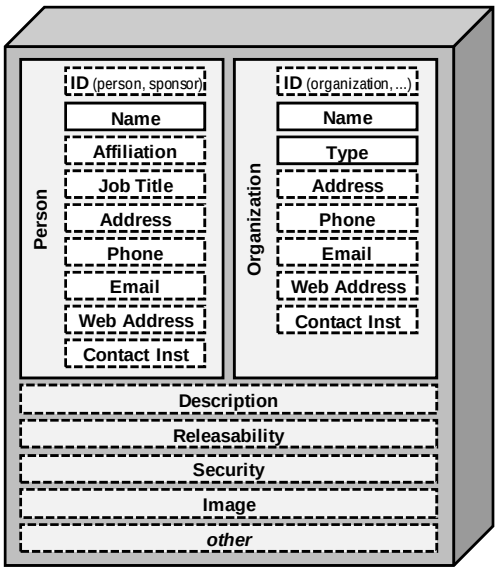

Figure 3: Contact Metacard Structure

Taxonomy Metacards

Many M&S resources refer to a specific taxonomy of terms and definitions. Beginning with version 1.4, the MSC-DMS provides a means to capture M&S taxonomies in a way that allows the discovery metadata relevant to a resource to be understood more richly. An M&S taxonomy metacard identifies a glossary of terms used by an organization, program, or focus group that can be attributed to one or more M&S assets. Such taxonomies provide a greater context for understanding the underlying metadata descriptions provided within an M&S resource. An M&S resource metacard may include citations to relevant taxonomies that have been defined.

Figure 4 provides an illustration of the MSC-DMS taxonomy metacard structure. The boxes with a solid outline represent required metadata components, whereas the boxes with a dashed outline represent optional metadata components.

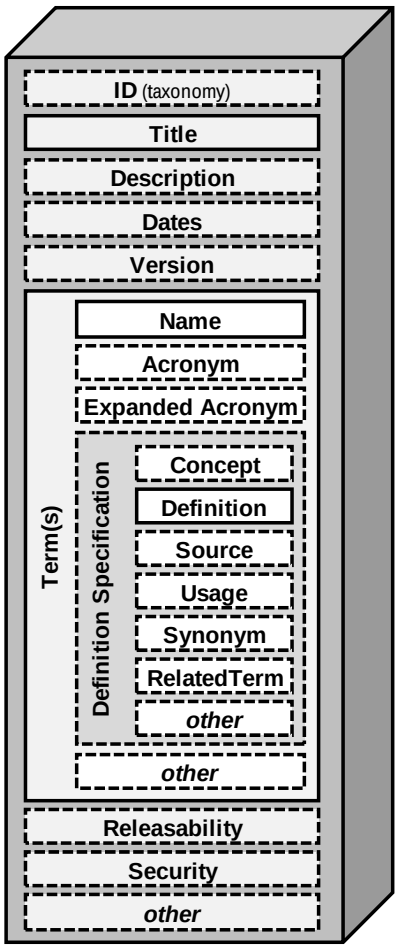

Figure 4: Taxonomy Metacard Structure

The taxonomy cited by a resource metacard is intended to reference a taxonomy document that conforms to the template provided by this structure.

Multicards

Many developers, integrators, and their organizations need to be able to package and distribute multiple sets of metacards, not just a single metacard. The MSC-DMS provides a means to package multiple metacards including resources, contacts, and taxonomies using a multicard.

Figure 5 provides an illustration of an M&S multicard structure. The boxes with a solid outline represent required metadata components, whereas the boxes with a dashed outline represent optional metadata components.

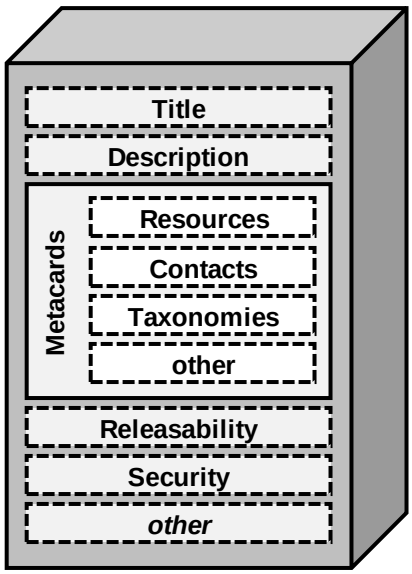

Figure 5: Multicard Structure

METACARD TOOLS

Metacard tools are emerging that help users create metacards for describing assets of interest, and provide a means to collaborate, validate, and submit metacards to repositories and catalogs. The chief goals for these metacard tools are the following:

- To make it easy to build and share M&S asset metacards
  - No need to be an XML expert
  - Automatic compliance of available information with the MSC-DMS standard
- To help facilitate M&S resource repositories (MSRRs) and the M&S Catalog to reflect more assets of interest

These emerging tools help ease the burden in managing metacards, in generating conformant XML-formatted metadata, and in submitting the metadata to the M&S Catalog (or any other authorized subscriber). Potential data sources that need this type of capability for generating M&S metacards are encouraged to contact the authors or M&S CO [12].

Metacard Builder

Metacard Builder is a Java-based tool that allows Metacards to be quickly built and edited without the need for understanding XML or the underlying XML schema. XML is the primary format used for most metacards including the MSC-DMS. Building XML documents can be an arduous and sometimes a complex process even for those who are XML savvy. Metacard Builder is intended to help remove that barrier by providing a framework for filling-in the metacard information based on the MSC-DMS. It guides the end-user in filling out the information.

Figure 6 provides a screen shot of this tool. The form on the left hand side provides the input area for most users. An XML representation is also provided in an editable *Text View* panel on the right side. If changes are made in the form window, then those changes are reflected real-time in the *Text View*.

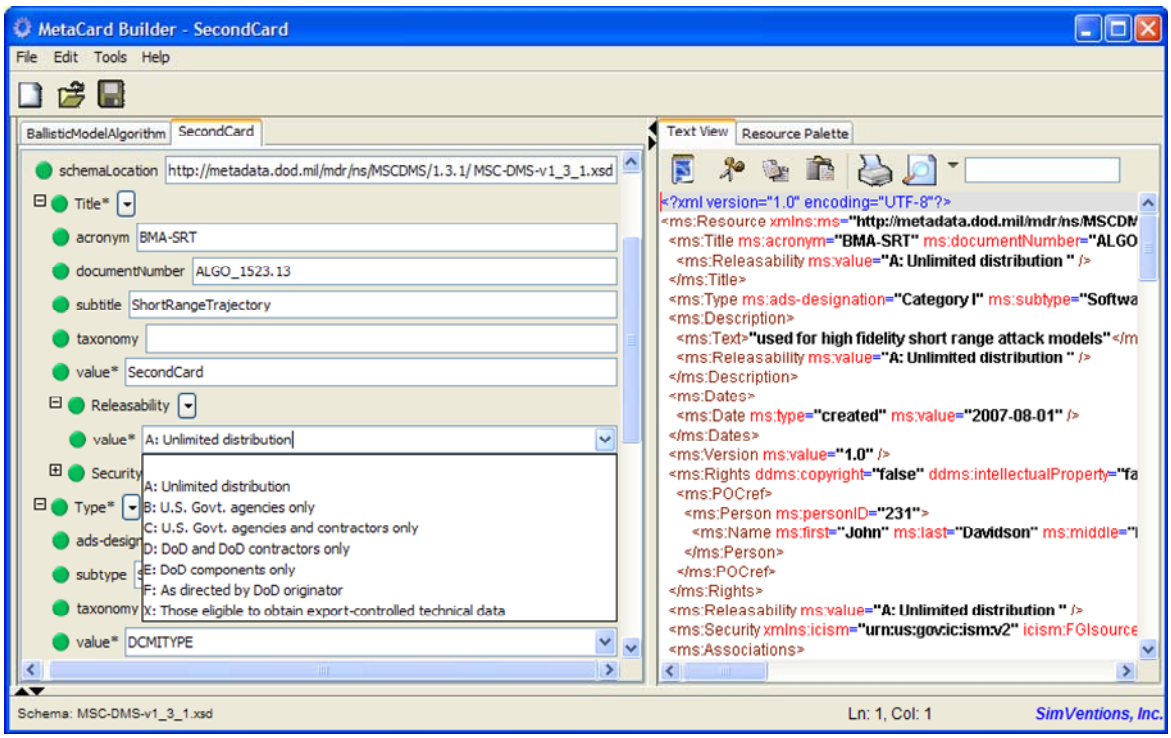

Figure 6: Metacard Builder

Some of the capabilities provided by Metacard Builder include the following:

- User-friendly, form-based management of metadata (e.g., enumerations in dropdown boxes)
- Validate metacards against user-selected schema
- Save and reuse frequently entered information
- View generated XML as user builds the metacards
- Support for multiple POCs and taxonomies
- Transformation of metacards to other schemas / data formats
- Search locally for metacards
- Support for multi-resource metacards (multicards)

EMBR Portal

The Enterprise Metacard Builder Resource (EMBR) Portal is a recently developed web-based environment that serves as a DoD M&S Catalog Interface Tool. A screen shot of the EMBR Portal user interface is provided in Figure 7.

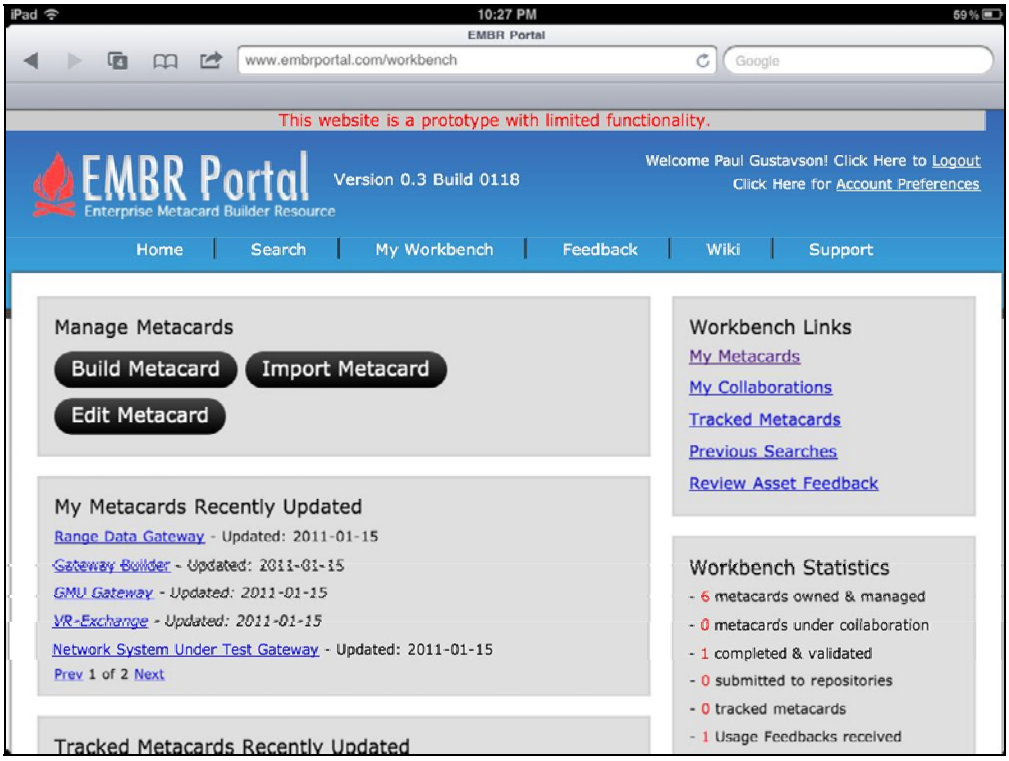

Figure 7: EMBR Portal

The EMBR Portal has been sponsored by M&S CO in support of advancing the LVC Architecture Roadmap (LVCAR) initiatives. The capabilities of Metacard Builder are contained within EMBR Portal, thereby providing a web-based framework that aligns with some of the core principles of cloud-computing. This includes data storage, security, and services.

The current EMBR Portal prototype, which enables the discovery, management, deployment, and transformation of resource metacards, is available for use to the DoD and approved contractors at https://embrportal.com. It specifically uses the MSC-DMS as the principal mechanism for documenting and storing resource metacards.

Operationally, the EMBR Portal works with current release versions of most web browsers including Microsoft Internet Explorer, Mozilla Firefox, Apple Safari, and Google Chrome. Additionally, the EMBR Portal operates on a server hosted on a Linux system with Apache Tomcat with access to a Solr search indexing platform and Lucene search library. Finally, the EMBR Portal is designed to permit access via the Internet only for those users who have secured an authorized account.

The specific product features that are provided by the EMBR Portal to support the development, discovery, and management of M&S asset metacards are identified below:

- Login and registration capability to help maintain a secure environment
- Search and discover capability to help potential consumers and integrators locate candidate M&S asset metacards that have been previously uploaded or built within the EMBR Portal
- Comment on existing metacards
- Bookmark capability to track metacards of interest

- Personal workbench capability that provides a listing of previous search queries, and a listing of bookmarked metacards
- Access to a personal data locker that contains metacards managed by the individual
- Ability to request collaboration with another metacard owner
- Ability to edit or build metacards
- Feedback mechanism to provide asset usage experience within an M&S asset metacard
- Ability to monitor usage experience feedback
- Collaboration capability providing a means for two or more people to work on and discuss development of a metacard
- Online support system to enable users to submit trouble tickets and seek assistance on using the EMBR Portal
- External interface, provided via web services, that enables the Portal to exchange data with existing catalogs, repositories, and registries

The EMBR Portal simplifies things by not only providing Metacard Builder-like capability accessible via the web, but also in providing a mechanism for producers, integrators, and consumers to manage their metacards of interest within a secure data locker, which is identified as the EMBR Workbench.

The EMBR Portal functionality, as it relates to the common stakeholders identified by the net-centric data strategy, is summarized in Table 1.

Table 1: EMBR Portal Functionality

	Producer	Integrator	Consumer
Create Metacards	Yes		
Edit/Collaborate on Metacards	Yes	Yes	
Search for Metacards	Yes	Yes	Yes
Post Metacard Feedback		Yes	Yes
Evaluate Metacard Feedback	Yes	Yes	Yes
Post to Community Page	Yes	Yes	Yes

The six functional capabilities are intended to help augment other systems including the M&S Catalog and various MSRRs. For example the ability to create, edit and collaborate on a metacard results in a validated and updated metacard that can be submitted to the M&S Catalog.

Another view of the capabilities across the three types of stakeholders is identified in Table 2 below.

Table 2: Stakeholder Capabilities in EMBR

Capability	Developer	Integrator	Consumer
To develop / modify and metacard	create / update	update / augment	augment (user experience)
To discover metacard instances	see how it has proliferated	proliferation / availability	see what's available
To evaluate application use	gather feedback (for update)	feedback / learn	learn how it's been used

The EMBR Portal continues to evolve and improve with a clear desire to move from a prototype to a production capability, which can be deployed in several ways. The vision is that it can be used as a single secure web site on the Internet with access from and submission to the M&S Catalog and other repositories, or it can potentially be deployed as a capability within an organization's or company's intranet environment, providing a common shared tool to build and manage metacards for that organization.

The EMBR Portal development continues to focus on providing a set of intuitive user controls (using browser standards) and web service interfaces for enabling management, deployment, and transformation of metacards. Figure 8 highlights the overall capabilities intended to be provided by the production version of the EMBR Portal.

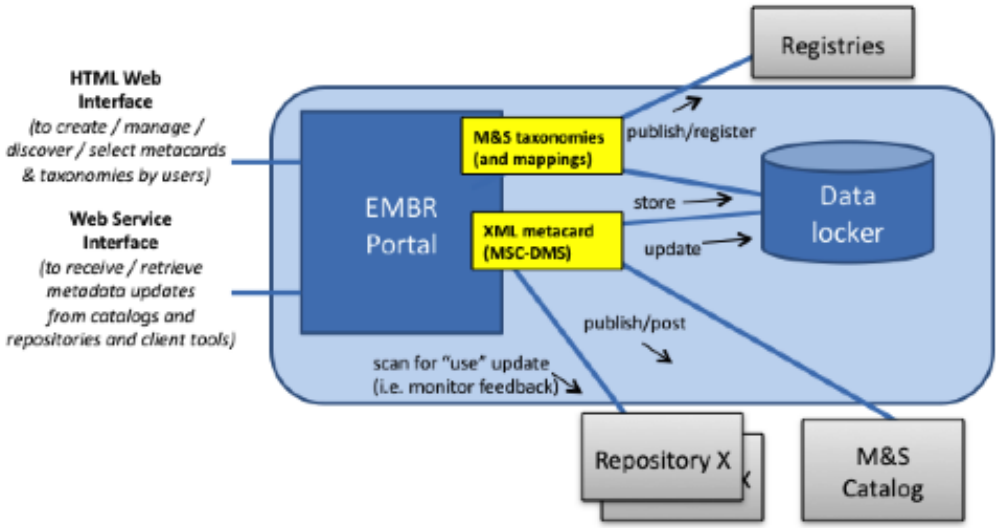

Figure 8: EMBR Portal Functionality

While much of the functionality can be demonstrated within the current prototype, the ability to submit to various repositories and a greater set of web sites is still being developed. Additionally, future updates of the EMBR Portal are expected to migrate from version 1.3.1 of the MSC-DMS to version 1.4 of the MSC-DMS thereby allowing not only M&S resource metacards to be built and managed, but also M&S contact metacards and M&S taxonomy metacards. Furthermore, the EMBR Portal is designed to potentially allow other types of metacards to be built and managed beyond the metacards defined by the MSC-DMS.

Future versions will focus on additional support in the following areas: metacard management, access control, improved search capability including use of taxonomy matching, increased social networking for sharing ideas, and improved site/repository integration with other sites and repositories.

THE DoD M&S CATALOG

The DoD M&S Catalog provides a common portal with connection to various M&S related repository sources. It achieved initial operational capability in 2010. Figure 9 illustrates end-to-end data flow from a source to the M&S Catalog. Each source has its own way of storing data. To communicate, the source must translate that data into a format understood by others. The metadata generation tools discussed above and the MSC-DMS are designed to serve that function.

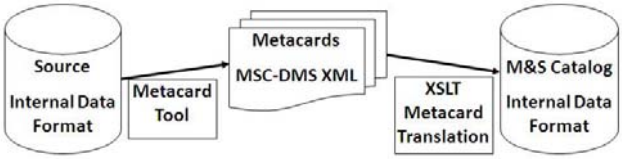

Figure 9: Data Flow from Source to M&S Catalog

Internally, an Extensible Stylesheet Language Transformations (XSLT) script translates from the MSC-DMS into the commercial off-the-shelf (COTS) format native to the M&S Catalog implementation.

M&S Catalog Sources

The following sources currently supply M&S metacard information to the M&S Catalog. More are in the works.

- Army MSRR (Army Modeling & Simulation Resource Repository)
- AFAMS (Air Force Agency for Modeling & Simulation)
- Navy MSRR (Navy Modeling & Simulation Resource Repository)
- AFSPC (Air Force Space Command)
- AFOTEC (Air Force Operational Test & Evaluation Center)
- ADTC (Army Digital Terrain Catalog)
- CAPE (Cost Analysis and Program Evaluation)
- DVDT (DoD VV&A Documentation Tool)
- IW (Irregular Warfare) M&S Senior Coordinating Group

- JMETC TENA (Joint Mission Environment Test Capability Test and Training Enabling Architecture)
- LVC-AF Repository (JFCOM Live/Virtual/Constructive Architecture Framework)
- MFS (NAWC AD Manned Flight Simulator)
- MSIAC (DoD Modeling & Simulation Information Analysis Center)
- NASA JEOD (Johnson Space Center Engineering Orbital Dynamics)
- Navy VV&A (Verification, Validation, and Accreditation) Document Archive
- NPS MOVES (Naval Postgraduate School Modeling, Virtual Environments, and Simulation Institute)
- OAML (Oceanographic and Atmospheric Master Library)
- PEO IWS (Program Executive Office Integrated Warfare Systems)
- RDG (Rapid Data Generation)

There are strong indications that the number of M&S contributors and customers for the repositories will continue to increase, and that the quantity of metacards from current contributors will continue to increase as well.

M&S Catalog System Architecture

Figure 10 illustrates the system architecture of the M&S Catalog.

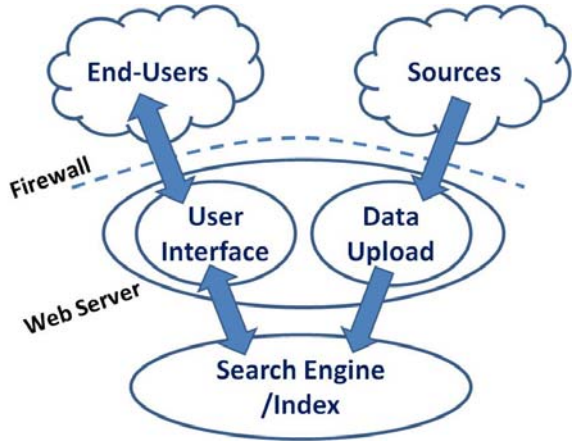

Figure 10: M&S Catalog System Architecture

For security reasons, the M&S Catalog itself is protected from the Internet at large by a firewall. Specific ports and protocols are held open to allow end-users (with the necessary DoD-issued or DoD-approved security certificates) to interact with the user interface located on a web server. A query received from an end-user is passed from the web server to the search engine. A response from the search engine is sent back to the web server, which formats it neatly and presents it to the end-user. Similarly, sources submit metacards to a data upload service on the web server. After automated validation and review, the metacards are passed to the search engine. The search engine then adds the new metacards to its indexes, from which it will serve future queries. Overall, the Catalog architecture shields the search engine and index from direct contact with the outside world but enables the transactions necessary to conduct business.

How the M&S Catalog Works (User Perspective)

In 2010, the M&S Catalog project took advantage of the lessons learned using an in-house Government off-the-shelf (GOTS) query interface tool to acquire and configure a commercial search engine to better meet the requirements. The current M&S Catalog engine is available for use to the DoD and its contractors at https://mscatalog.osd.mil. One of the capabilities is a "faceted search," in which the search engine will provide the user with a succession of search options. One well-known example of this is seen in the typical merchant website that offers filters based on product type, brand name, capacity, and price. The M&S Catalog actually provides *dynamic* faceted search, in which options are presented based on the underlying data currently selected, and are refined as the search narrows.

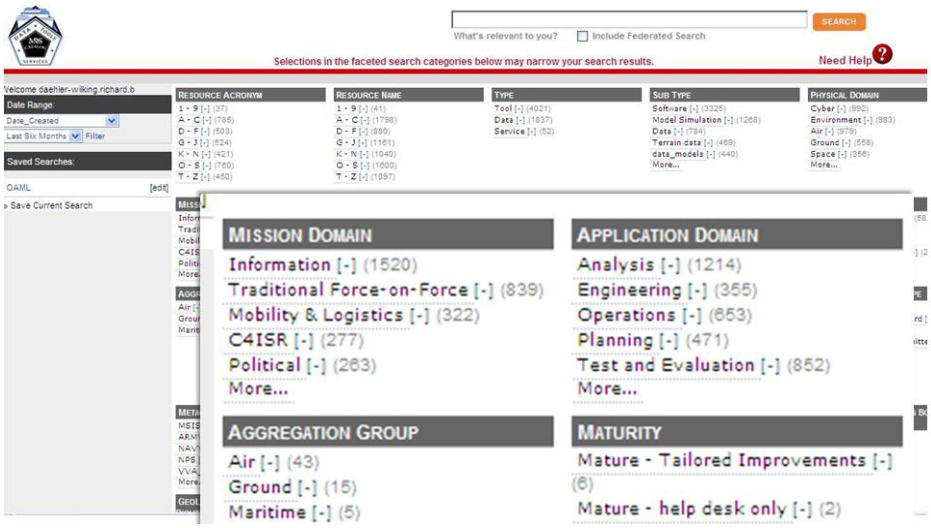

Figure 11: M&S Catalog Initial Search Screen

A specific example of dynamic faceted search is related to the keywords structure of the MSC-DMS. This allows for ordered pairs consisting of taxonomy and value. These are expressed in the M&S Catalog as facets. For example, some records specify mission domain, with values like C4ISR, information, irregular warfare, mobility & logistics, etc. (see Figure 11). Other records specify functionality, with values like adaptive planning, agent-based, algorithmic tool, analysis, etc. (not shown here). The user is allowed to select a value from any such facet, and the selections then offered to the user vary according to what has been selected. Thus, if the user chooses mission domain = irregular warfare, the choices for other facets are reduced. This is because only the reduced collections of facet values are represented among the records associated with the mission domain of irregular warfare.

On the other hand, filtering for a mission domain of irregular warfare does not eliminate the mission domain facet. As seen in Figure 12, it is still possible to select a few other values. This is because one record may have multiple mission domain values. In fact, Figure 12 shows us that 14 records also have a mission domain of information; one has mobility & logistics, and so on.

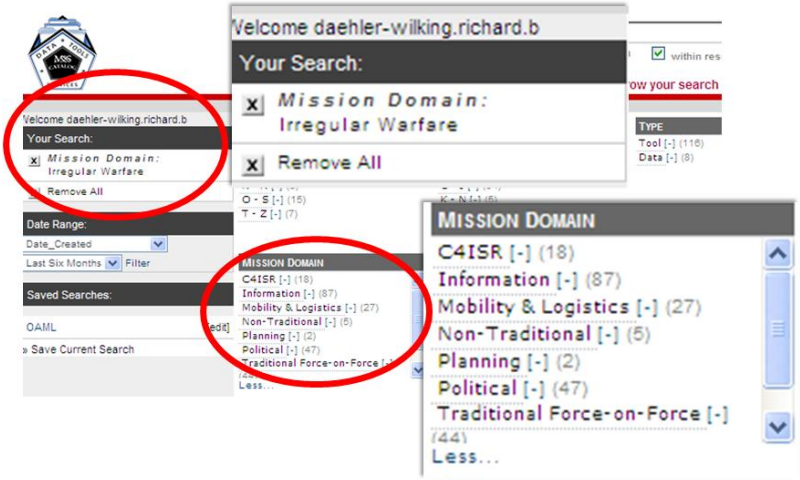

Figure 12: Filtered for Irregular Warfare

Making such searches possible requires some advance preparation. It is necessary to capture keyword information in a record before it can be used by the M&S Catalog. To be practical, this requires all sources to use the same set of values for a particular facet, such as "irregular warfare" (spelled out) rather than "IW", or "agent-based" (hyphenated) rather than "agent based" (no hyphen). This boils down to the use of common taxonomies, as supported by v1.4 of the MSC-DMS.

Of course, all of this assumes that the M&S Catalog has metadata records to be searched. A key priority in 2011 and beyond continues to be to sign up more sources and sources of different kinds of data supporting experimentation, training, analysis, and testing -- all the different DoD communities enabled by M&S.

When a source does sign up for the M&S Catalog, it submits metacards describing its resources. Ideally, this is in the form of XML files that comply with the MSC-DMS schema, but we have learned from experience that it is beneficial to meet each source where it is now, and try to help it grow from there. The topic of metadata generation tools to assist sources is covered separately above. It is also addressed in documentation available at the M&S Catalog website [13] as indicated in the left margin of Figure 12.

In whatever form the source produces its metadata files, they can contain one record each or can contain an aggregate of multiple records. The files are submitted by each source using either e-mail or an upload web-service hosted by the M&S Catalog. If the source submits an XML file based on MSC-DMS, the file is scanned for compliance with the schema and feedback (acknowledgment or error message) is returned to the submitting source for each individual record, as illustrated in Figure 13.

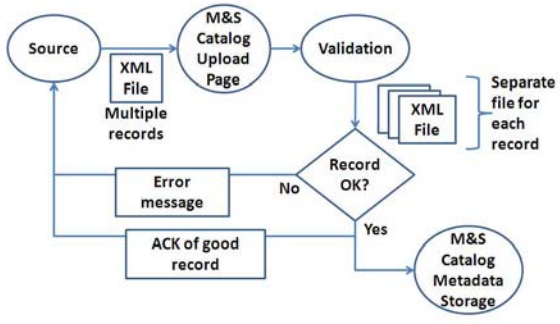

Figure 13: M&S Catalog Upload Work Flow for XML Files

The source provides metadata corresponding to the MSC-DMS data elements, using as many of the MSC-DMS elements as applicable to the resource being described. At a minimum the metacard data provided by each source must include:

- Resource ID
- Title
- Description
- POC

Certain other fields are highly recommended but not absolutely required for source submission to the M&S Catalog. To be compliant with the MSC-DMS specification, the following fields must also be present:

- Type
- Dates
- Version
- Releasability
- Keyword

Very important is the fact that the M&S Catalog explicitly wants to stay as closely aligned with the MSC-DMS as possible. To date, three factors have contributed to the differences between the M&S Catalog's data model and the MSC-DMS.

1. The M&S Catalog is on the "cutting edge" of defining metadata actually needed and used by the sources and end-users.
2. We have learned from experience that most sources are not yet able to generate MSC-DMS files, and will need something like the metadata generation tools discussed above to do so. The M&S Catalog has come up with some tools of its own, but they have perforce been tied primarily to the M&S Catalog's own "practical" data model rather than the broader MSC-DMS.
3. The M&S Catalog is implemented in a COTS tool that emphasizes quality of the end-user's search experience, rather than support of complex XML data schemas. The structures that support fast and flexible dynamic facetted search as described above do not allow nested data or data of arbitrary values. For instance, it is practical to set up a facetted search for a particular U.S. state because there are only fifty states to choose from (plus the District of Columbia and a few other places). It is even possible to set up search for ZIP code ranges (e.g., something like 29xxx for South Carolina). However, it is not at all practical to set up a facetted search for all street names – there are simply too many of them to list. Therefore, the

address of an entity cannot be entered as a facet. The only other choice is as a block of text data. If it is to have structure (street/city/state/zip), that structure must be handled by internal formatting and custom-written code in the user interface; but it is not natively supported by the search engine itself or by the input file format accepted by the search engine.

Nonetheless, the vision and process are to keep the MSC-DMS and the M&S Catalog going in the same direction. Thus, all data elements in the M&S Catalog's data model can be mapped to the MSC-DMS (e.g., facets to keywords). When something is encountered that does not map well, the M&S Catalog project works with the MSC-DMS team to align things as well as possible, then submits a recommendation for any changes needed in the MSC-DMS.

DoD DATA DISCOVERY RELATED WORKING GROUPS

In the previous section we saw that metadata about M&S assets can be generated and transmitted in a variety of ways. We are also discussing the differences between the current implementation of the M&S Catalog (focused on the end-user's ability to find things quickly and easily) and the more general need to store information in a structured way (so that the information can be used after it has been found). There is also focus on what is most easily used by consumers and what is most easily created and maintained by producers.

These topics all come into play as DoD communities enabled by M&S strive to establish mechanisms like the MSC-DMS that serve the needs of all concerned. To address these needs, M&S COI has established DoD Data Management Working Group, which includes the Discovery Metadata Team.

SUMMARY

The ability to discover existing M&S assets is a critical need for enabling effective reuse and for reducing the duplication of capabilities in the development and selection of appropriate M&S assets. In the last three years, the MSC-DMS, the emergence of metacard tools, and the M&S Catalog, including its connection to various repository sources, have made it easier for DoD organizations to build, connect, discover, and share M&S resources. However, this is only the beginning. The DoD M&S community is encouraged to continue to leverage the MSC-DMS and supporting metacard tools in creating well-described and easily discoverable M&S resources and to utilize the M&S Catalog as the way to query and locate M&S resources of interest. We seek new sources to add to the collection of connected resources available to the M&S Catalog. Furthermore, the M&S community is encouraged to share their experiences and provide their feedback so that the MSC-DMS, the M&S Catalog, and associated tools such as the EMBR Portal can be better understood and further improved.

In this article we've highlighted the importance of discovery metadata. We have shown that metadata drives visibility, accessibility, and understanding as identified by the DoD Net-Centric Data Strategy. We've explored how the MSC-DMS provides a standard for the DoD M&S community to catalog metacards for three types of M&S assets: resources, contacts, and taxonomy glossaries. Specifically we examined how resource metacards provide a consistent labeling convention for simulations, federations, adjunct tools, data models, interface models, and related documents. These M&S resource metacards, based on the MSC-DMS, are what hold the key information for describing, locating, and using M&S resources.

We have also highlighted a principal mechanism to discover assets of interest through the M&S Catalog, which provides an overarching discovery tool. The M&S Catalog, the metacards, and the metacard tools and repositories are illustrated in Figure 14.

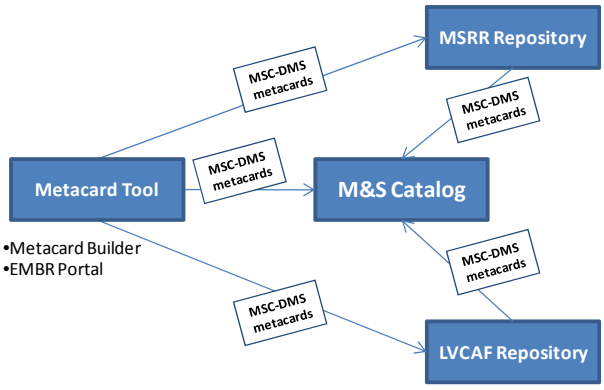

Figure 14: Metacard Data Flow from a Source

However, all these things are dependent on the data sources, data proponents, data curators, and repositories to document these assets and share them through the M&S Catalog. The need of the communities influenced by M&S should be unmistakable – to make data visible, accessible, and understandable. This starts with the producers, then

integrators, and finally feedback provided by consumers – well-documented information that drives decisions.

ACKNOWLEDGEMENTS

The authors wish to acknowledge M&SCO, JTIEC, and the DoD M&S user community for their contributions and suggestions in helping refine the characteristics of the MSC-DMS, supporting tools, and the M&S Catalog. Additionally credit is also due JHU/APL for their foresight in leading the EMBR Portal effort in response to LVC Interoperability Reuse needs.

REFERENCES

[1] Gustavson, Daehler-Wilking, Blais, Rutherford, Dumanoir, "Discovery and Reuse of Modeling and Simulation Assets", 10S-SIW-048

[2] Gustavson, Nicolai, Scrudder, Maximizing Discovery of M&S Resources -An Inside Look at the M&S COI Discovery Metadata Specification (MSC-DMS), 08F-SIW-043

[3] Gustavson, Nicolai, Scrudder, Blais, Daehler-Wilking, "Discovery and Reuse of Modeling and Simulation Assets", 09S-SIW-076

[4] The illustration of the parts of a food label is from the New Jersey Department of Health and Senior Services (DHSS) Food and Drug Safety Program. It is available online at http://www.state.nj.us/health/foodanddrugsafety/documents/parts_of_food_label.pdf

[5] The Oceanographic and Atmospheric Master Library Summary (U), Naval Oceanographic Office, Systems Integration Division, Stennis Space Center, MS 39522-5001, Document OAML-SUM-21S, April 2010. This document is unclassified, but subject to Distribution Statement C: Distribution authorized to U.S. government agencies and their U.S. contractors. Requests for products are to be submitted to the OAML Secretary at walter.moskal@navy.mil .

[6] Department of Defense (DoD) Metadata Registry And Clearinghouse, https://metadata.dod.mil/mdr

[7] Modeling and Simulation (M&S) Community of Interest (COI) Discovery Metadata Specification (MSC-DMS), DoD M&S Coordination Office, Preliminary Release, Version 1.4, 24 December 2010.

[8] DTIC, "Net-Centric Environment Joint Functional Concept," Version 1.0, April 7, 2005, http://www.dtic.mil/futurejointwarfare/concepts/netcentric_jfc.pdf

[9] Deputy Assistant Secretary of Defense, Department of Defense Discovery Metadata Specification (DDMS), Version 1.4.1, August 10, 2007.

[10] The M&S Glossary was originally published in 1998 and is available at http://www.msco.mil/files/MSCO%20Online%20Library/DoD%205000.59-M%20-%20MS%20Glossary1%20-%2019980115.pdf. An updated version is expected soon.

[11] MSC-DMS Implementation Guide, DoD M&S Coordination Office, Version 1.4, 24 December 2010.

[12] The MSC-DMS schema and its implementation guide are available online at the Department of Defense (DoD) Modeling and Simulation Coordination Office (M&S CO), http://www.msco.mil/resource_discovery.html

[13] The DoD M &S Catalog is accessible to those with DoD or DoD-approved security certificates at https://mscatalog.osd.mil

[14] Riggs, Morse, Brunton, Gustavson, Rutherford, Chase, Belcher, "Emerging Solutions for LVC Asset Reuse", 11S-SIW-034