2005 Spring Simulation Interoperability Workshop
San Diego, California, April 2005

Web Services based on the C2IEDM - Data Mediation and Data Storage

Dr. Andreas Tolk, Saikou Diallo, Kevin Dupigny, Bo Sun, Chuck Turnitsa

Virginia Modeling Analysis & Simulation Center (VMASC)
Old Dominion University
Norfolk, VA 23529
atolk@odu.edu, sdiallo@odu.edu, kdup0001@odu.edu, bsun@odu.edu, cturnits@odu.edu

Keywords:

Service Oriented Architectures, Multi-Resolution Modeling, Web Services, Web Service Interest Management

ABSTRACT: *This paper shows how the ideas of web services and Model-Based Data Management (MBDM) are merged in two "C2IEDM Web Service" families realized by the VMASC team within the Battle Management Language (BML) group. Although being developed within the team for the prototype, the services can be used in other systems and projects easily, as they can be adapted and configured to serve alternative needs.*

To support unambiguous definition of data element for information exchange, the ideas of MBDM have been implemented, resulting in a data mediation service based on the use of the Command and Control Information Exchange Data Model (C2IEDM) as a common reference model. Generally, data engineering includes four parts: data administration, data management, data alignment, and data transformation, of which data management, alignment, and transformation are of particular interest within the BML group. Applying MBDM, results are documented for data alignment using a commercial tool to map, aggregate, and disaggregate data, and remove/create duplicated data. This results in software products and configuration, that are directly used for data transformation: The commercial solution supports XSLT and Java-based mapping methods to set up transformation layers between both XML and database approaches. In BML, this is used to set up a C2IEDM based data mediation service speaking C2IEDM as well as the language to be mediated (in our case BML).

The second service family contains C2IEDM based information exchange and storage services, which use the tag set of the coalition name space of the US DoD XML Repository to store, retrieve, and exchange information based on the C2IEDM. These services can be source and target of simple C2IEDM data replications. When combined with the C2IEDM based data mediation service, the C2IEDM based information exchange, and storage service can be used to store and exchange information of every mediated language.

The two service families can be used in M&S infrastructures, such as Extensible M&S Framework (XMSF) prototypes, the Joint National Training Capability (JNTC), or the Distributed Continuous Experimentation Environment (DCEE), as well as in Global Information grid (GIG) infrastructures, hence, connecting both worlds effectively and efficiently.

1 Introduction

The work described in this paper has its academic roots in many articles presented to the Simulation Interoperability Standards Organization (SISO) before, such as the following:

- The first recommendation to use the Data Model of the Allied/Army Tactical Command and Control System (ATCCIS) as a common information exchange data model was presented to SISO in 1999 in [1].
- Presenting the possible application fields of heterogeneous data federations were presented in paper [2].
- Recommendations to use the Command and Control Information Exchange Data Model as done in this paper have been summarized in [3].

The principle ideas of data engineering in general were also presented to a broader audience in [4]; this generated very positive and encouraging collaborations on related issues contributing to the work described in the following sections.

While many of these papers were more or less theory, this paper presents how the ideas were applied within the portions of work conducted by the Virginia Modeling Analyses & Simulation Center (VMASC) related to the projects Extensible Battle Management Language (XBML) and Air Operations Battle Management Language (AO BML). In both projects, VMASC was integrated into an overarching team comprising partners of Alion Science & Technology, Atlantic Consulting Services, Inc., George Mason University, and Gestalt, LLC. The work described in this paper is therefore just a small contribution to the general approach to couple Command and Control (C2) systems and M&S system based on a common Battle Management Language. Papers [5, 6, 7] will give the broader view on the ideas framing the work and go beyond the ideas described here. In particular, the mapping work done by our partners of ACS must explicitly be mentioned, as they are the basis of our approach [8].

Finally, the techniques described in this paper have alternative uses, specifically for developing prototypes in XBML and AO BML. We do see the role of VMASC as a generator of ideas and prototypical solutions for the future, while our industry partners focus on implementations which are operationally "safe enough for real users" in experimentations outside the "academic ivory tower." However, this does not mean that the implementation ideas presented here are not mature. We were able to integrate our services into the prototype of our partners within two days and did not have a single crash during the demonstrations during I/ITSEC 2004; however, our focus is to show what will be possible soon, not to create systems. The fact that VMASC is publishing these ideas separately in this paper is "academic culture"; it should not be seen as an attempt to decrease the value of the work of our industry partners. We are working in an emerging research & development area and alternative implementations and discussion of pros and cons is necessary for scientific progress; and it is this progress we want to contribute to with the ideas discussed in later sections.

The idea VMASC wanted to implement within the XBML and the AO BML project is described in detail in [3,4]:

We want to use an open standard web based approach to couple legacy systems of the C2 and the M&S systems based on the international information exchange for military operations: the Command & Control Information Exchange Data Model C2IEDM!

The result is a web service architecture that can be used to translate any XML schema into an equivalent C2IEDM based XML schema, if the information elements are sufficiently similar. This is done with the first web service family. Furthermore, XML documents supporting C2IEDM – either natively or after transformations based on the first web service set – can be used to populate a C2IEDM database implemented on open standards. This database is accessible via the second web service family. Finally, the information can be displayed using open web based approaches as well, again using web service to allow the use of alternative implementations, such as already existing C2IEDM implementations.

The authors see both service families as a possible contribution to the current ideas discussed within the NATO M&S Exploratory Team ET-016 on Coalition Battle Management Languages as well as for the SISO Study Group dealing with the same issue. The core idea can be summarized as follows:

- The works described in [2, 3, 4, 6, 8] show that information elements describing the same domain can be principally mapped to each other. Data engineering [4] defines a method.
- The military domain is sufficiently covered by the information exchange data elements as currently defined and described by the C2IEDM. A complete definition of every data element is part of the documentation.[1] The extensibility of C2IEDM to meet national concerns has been proven several times within NATO.
- Information exchange requirements can be modeled and documented in XML. Setting up XML interfaces for applications supporting information export and import is a standard procedure being tool supported.
- Our first service family translates an XML dialect into C2IEDM structured XML documents. As an application with an XML interface doesn't have to change – and the

[1] The complete documentation including scripts is available at http://www.mip-site.org

migration to XML is commercially supported by tools – this approach is a convenient way to make a system C2IEDM compliant.

- This C2IEDM compliant XML documents can be used either for immediate information exchange – as the translation goes both ways – or to store the information in a C2IEDM compliant database. This database can become source and/or target of data replication mechanisms as implemented for C2 systems, in particular of the newer generation.

These ideas are summarized in the conceptual schema depicted in Figure 1.

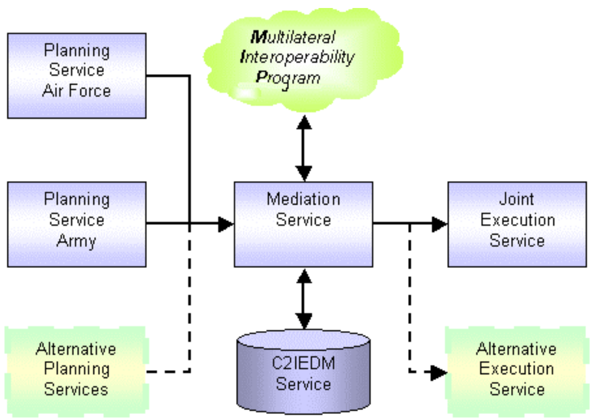

Figure 1: Battle Management Language Services

The necessary details will be discussed extensively in the following sections. VMASC is currently working on a version that can be downloaded for projects of interest to the sponsors, i.e., the Defense Modeling & Simulation Office (DMSO) and the Joint Forces Command (JFCOM), which will include international collaboration.

2 BML Service Architecture

Within this section, we will set the stage by describing the overall architecture. As stated earlier, the system is defined as one providing two different families of services, namely data storage (to the C2IEDM), and data mediation (between existing systems – simulation systems, C2 systems, etc). Before diving into a detailed description of the architecture and its various components, let us first look at some definitions.

A system's architecture is defined as *"the structure or structures of the system, which comprise software elements, the externally visible properties of those elements, and the relationships among them."*[9] Based on this definition, our architecture must be comprised of the elements that make up its intended processes (i.e. – data storage and data mediation), as well as the external (connectivity supporting) properties of those elements.

In order to accomplish this task, we chose to design the system using the principles of service-oriented architecture (SOA). A good working definition for SOA is a system for linking resources on demand. In an SOA, resources are made available to other participants in the network as independent services that are accessed in a standardized way. This provides for more flexible loose coupling of resources than in traditional systems architectures.[2]

In light of this definition, the use of an SOA to implement our systems becomes obvious. Furthermore, there are many other advantages to adopting this architecture, namely

- The state of the client does not affect the state of the service. The client (C2 system, Simulation System etc…) requests information in a clearly defined format through an interface, and the service provides the information in a fashion unknown to the client
- The disappearance of a client does not affect the overall system (which is not true for traditional architectures). Other clients continue to be served as needed. This means that only the service is aware of all of the participants thus making it the only mediator between the participating systems.
- Data mapping and data storage can be offered as distinctive components of the same service. This allows a substantial reduction of overhead data during the information exchange since static/setup data will only be communicated once.
- The SOA approach reduces the role of the client (subscribing systems) and puts the onus on the service to internally perform all the tasks and only return pertinent results

The most common implementation of SOAs is found in commercial web services. Web services are

[2] See http://looselycoupled.com/glossary/

standardized and accepted for all targeted application domains.

A further defining aspect of our architecture is that it is based on the principles of Battle Management Language. BML is defined as *"the unambiguous language used to command and control forces and equipment conducting military operations and to provide for situational awareness and a shared, common operational picture."* [5]

One of the stated goals of our system is data mediation between widely different C2 systems and other systems. To support that goal an unambiguous language such as BML is exactly what is needed to provide interoperability. This is done by providing for a standardized way to address troops, define and describe all elements that can be expected to be found within the battle space, and provide for past, present, and future (proposed) views of the interactions of troops with the elements of that battle space.

In other words: *BML describes missions to be conducted and means assigned to accomplish this mission within spatial-temporal constraints*; or even simpler: BML comprises military orders in the actual context: **Who** is doing **What**, **Where**, and **When**, and **Why** is he doing this. In the net-centric world, this 5W principle is sufficient as the local commander decides on the use of his resources within the constraints on its own. However, based on earlier concepts not stressing this form of *Auftragstaktik* (responsibility of the local commander based on the commander's intent), the two fields of **How** he has to accomplish the mission, and **Which** resources he has to use are defined as well. While this 6W+H scheme seems to be inappropriate for new military tasks with real military decision makers, it is a valid approach for simulation systems with insufficient interior decision logic. If a Semi-Automated Forces (SAF) system does not support the "How" and "Which" decisions needed to derive situation-adequate orders from the commander's intent for the simulated entities, this information must be provided by the BML user.

However, let us get back to the technology discussion. The various elements of our services oriented BML architecture accomplish the goals of data storage and data mediation in two ways. First, they give a method of communicating and describing data for the externally exposed connections into the system. Second BML provides a method of communication between the service elements.

So far, this is an appropriate first set of definitions of (1) what system architecture is, (2) why we chose services oriented architecture, and (3) why we have encapsulated all this within BML. However, we did not describe yet what our individual services are and how they support each other and connect to outside systems.

As mentioned above, our system really entails two families of services. These two service families are data mediation, and data storage.

Within our architecture, data mediation is the web service that allows for the transfer of data between two (or more) distinct and different systems. This is accomplished through the standard web services practice of having an exposed web services definition language (WSDL) interface that describes access to the service. As intended, the WSDL gives a definite description of the data elements that are needed to access the service, and a definite description of what is returned, but it does not provide a description of what manipulations the service provides on the data. The underlying assumptions are not defined by the WSDL. To address this, at least in our system, we rely on the fact that BML is intended for unambiguous communication – meaning that the elements comprising the 5Ws are self-explanatory.

For systems to interoperate through our mediation service, it is necessary and sufficient that they communicate with the BML view that the WSDL presents. Additional systems can join via the same mechanism. This flexibility based on only one data translation is the key to the exceptional value that our data mediation service provides. If each system that wishes to communicate with any number of other systems takes care to ensure that its data communication adheres to the WSDL, it is then (with one translation of it's internal data to that WSDL) able to communicate with all other systems that are thus compliant.

The second service family provided for by our system is that of data storage. The data storage service connects internally to the data mediation service and pulls information from it; again with a service connection based on the principles of BML. The data storage service records data in a relational database identical to the C2IEDM data reference model. The storage service is treated, conceptually, as another outside system by the mediation service. This allows for any outside system compatible with the WSDL to also have its data stored within the C2IEDM enabled data storage service. The following figure illustrates how the separate systems communicate with each other through our service families, and how that communication enables the concept of direct system-to-system connectivity.

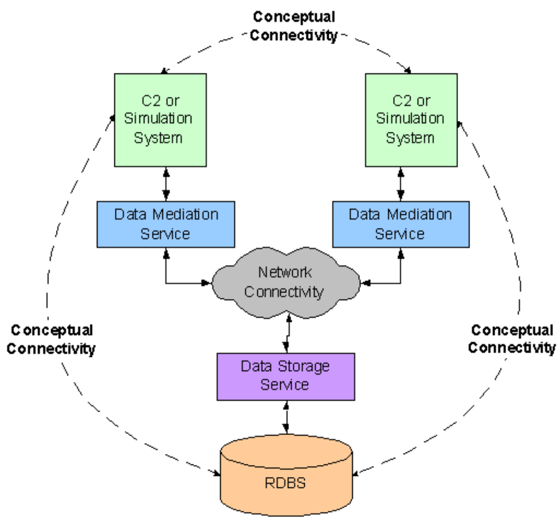

Figure 2: Conceptual View of the SOA

Currently, we are discussing

1) Whether the WSDL is sufficient as a "container" for data describing BML or is additional information needed for BML, such as ontology languages, and
2) Whether the structure of the XML schema supported by the WSDL should directly reflect C2IEDM data elements, or it would be better to use data structures directly derived from BML work, such as the 5W elements.

This discussion goes beyond the scope of this paper; it is conducted in detail in other papers. The approach discussed in this paper uses web services and C2IEDM elements, but the technology can be applied to other reference data models and can also be extended to ontology languages.

3 BML C2IEDM Services

Within this section, we will deal with the C2IEDM services having been defined and implemented for the BML project in more detail. We will start with the definition of the XML tag set used to exchange information, and then will present the two service families: data mediation and information exchange & storage services.

3.1 The C2IEDM XML Tag Set

XML is a well-formed text based markup language that is increasingly used as the standard for information exchange on the web as well as for data documentation. The purpose of XML is to describe, identify, and even qualify the data contained in a document. XML is used for traditional data processing, document-driven programming, archiving and binding. In addition, XML has the advantage of being hierarchical, linkable, easily reusable, and easily processed. It makes perfect sense and becomes common practice to use XML as a mean to describe database contents. In our work at VMASC, we use XML to support the implementation of the mapping from the BML data to the C2IEDM data model (data mediation) as well as to access the C2IEDM database (data storage).

The information generated from C2 systems is generally stored in relational databases. In order to take full advantage of the power of XML, the data needs to be represented in XML in accordance with the W3C standard for well-formed documents. Furthermore, many software packages can be used to directly move a set of data from a traditional relational database view to a more flexible XML view. Our evaluation of possible tools resulted in the selection of Altova MapForce [10], which supports different coupling methods, languages, and platforms, and which does not only support basic mappings but also complex mappings (see section 0 for details). It provides developers with the ability to connect to most of the relational database and generates a well-formed XML document. Furthermore, MapForce can create a document based on the raw data provided in tabular format for projects that do not require schemas. However, in addition, from the relational database it has the capability of creating an XML schema that maintains all relationships and dependencies of the source model, and respects identity constraints.

The tag set thus generated is directly related to the original data set. This method has the advantage of preserving all of the previously established data structures and their relationships. Therefore, there is no loss of information when the XML schema is generated. Furthermore, since data interoperability can be defined as the ability to reuse data from one C2 system to another without outside intervention, it is then possible to use available XML tools to map the contents of any number of data sources.

Providing a tag set directly related to the source database in the fashion described earlier in conjunction with the adherence to the DoD standards holds several key advantages. The most important one is the fact that it allows the implementation of the mapping solution as a service oriented architecture taking full advantage of the power and flexibility of commercial web services. It also allows the creation of a highly reusable and transportable data-mapping tool using XSLT. This tag set is used extensively to implement the data mediation service.

In addition to these more general observations, there is a particular advantage within the US: In order to align the name spaces within XML oriented applications within the US DoD, in particular within the Global Information Grid (GIG), the US DoD XML Registry was established. This repository of XML tag sets and schemas is used to collect all relevant XML tag sets used within the responsibility of the US DoD. In addition to the DoD XML Registry, where XML tag sets are simply registered, the U.S. Department of Defense established the "DoD Metadata Registry and Clearinghouse." Similar to the hierarchy of applicable doctrines, the XML Registry is structured as well into coalition namespaces, joint name spaces, service name spaces, etc. The first entry into the coalition namespace is an XML version of the C2IEDM derived by experts of the Institute for Defense Analysis (IDA). Using this tag set ensures that all advantages envisioned in papers like [3] can be automatically utilized.

3.2 The C2IEDM Information Exchange & Storage Services

As described in detail in [3], the Command & Control Information Exchange Data Model (C2IEDM) is an interoperability solution that was developed by NATO experts and which is now managed, improved, and distributed by the Multilateral Interoperability Program (MIP).

The utility of C2IEDM goes far beyond the battlefield for it can be used in live, virtual, or constructive training scenarios. The data model is a medium for information exchange and transmits the data that is needed or requested rather than all the data it houses. This specific data transfer feature allows multiple C2 systems to then become interoperable and pull only the data it needs from the data repository, which they share.

The C2IEDM standard specifies only a need for the C2 system conform to the information architecture leaving all other matters of operations to the connected systems. This naturally creates a hub-like architecture in which the C2IEDM Database repository rests in the middle of any set of C2 systems. The central information hub is an idea similar to the premise that is driving the current development GIG. Multiple nations, not limited to NATO partners, and various experts have contributed to the development of C2IEDM. Special care was taken in order to make the standard extendable, as information requirements of C2 systems are liable to change as time goes by. The C2IEDM contains three well-integrated views:

- The conceptual model within C2IEDM covers concepts with any remote military significance i.e. organizations, actions, reporting capabilities, land features, weather, etc.
- The second model is a logical model in which the concepts are linked together first forming aggregation relationships (i.e. organization, units), and then cross concept relationships (i.e. actions, units/organizations carrying out action).
- The final model is then the physical model, which allows C2IEDM to be represented as a physical database where actual data can be stored. This physical form can be represented as the database itself or entirely as an XML schema, as already explained in section 3.1.

The natural normalized structure of C2IEDM creates several areas of natural concept relationships. A mission statement is usually comprised of elements, which are important to make it complete and unambiguous. Following the terminology of the US Army, the BML Framework describes these elements as being the 5Ws (see discussions in section 2) of an order.

Exploiting C2IEDM's logical framework allowed the fast identification of a subset of concepts in which the 5Ws could be defined. Similar exercises are described in earlier papers referenced in [1,2]. The basic concept areas used for the 5W included:

- **Organization**, in particular the derived units, specifying the who
- **Action**, specifying the what and when
- **Location**, in particular in association with object items and actions, specifying the where

Each concept area is composed of multiple tables that enable linking and the explicit specification of information. Figure 3 shows the interpretation as used in the VMASC database. It should be pointed out that additional papers show the generalization of this idea and its extensions, in particular [8].

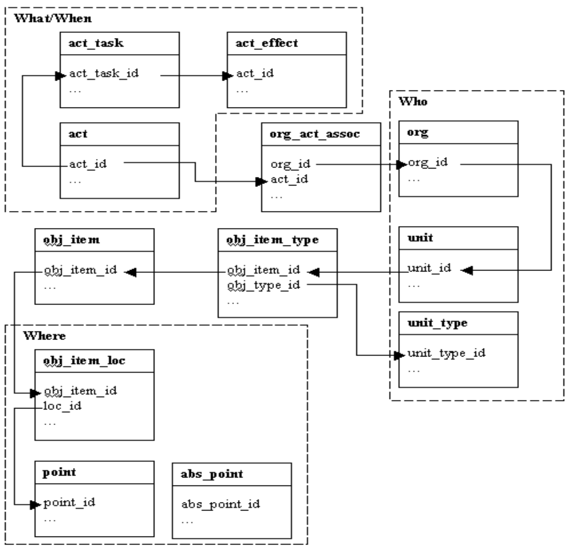

Figure 3: Sample Tables in C2IEDM Database

The storage service was implemented in form of a MySQL database. As the overall architecture uses web services to access the repository, XML interfaces based on the coalition XML namespace tag set were defined. The GUI of this database is implemented as a web service itself, and can be used independently from the database itself (only the service directly accesses the database).

For insertion purposes, the GUI presents pre-formatted templates to the user to insure the correct insertion of data. The C2IEDM format is very precise having a very small margin error. Each template once filled accesses the services, which allow them to fill the concept area of the database for which they are attached too. Section 4.1 describes ongoing work concerning the improvement of the GUI.

The retrieval of data requires a key in order to identify what or whose information is being requested. In addition to using system-defined keys, for testing and stand-alone purposes, the GUI has the feature to provide the user with a listing of all available organization keys that are in the system. These keys are used to access the organizations' information (a 5W mission). The retrieved information is displayed in a simple 5W format. As stated before, section 4.1 will show improvements.

In summary, this web service family implements a C2IEDM compliant MySQL database with an XML interface based on the C2IEDM XML tag set of the coalition namespace as submitted to the US DoD XML Repository. For access and display of the information, web services are utilized, so that display, access layers, and database can be distributed.

In addition, all components can be downloaded free from their distributors; all standards used are open and are easily obtainable.

3.3 The C2IEDM Data Mediation Service

We think that using C2IEDM to generate data mediation services utilizing the ideas and concepts of Model Based Data Management (MBDM) is a good example to support unambiguous definition of data element for information exchange [3]. As defined in [4], data engineering based on mediation services of C2IEDM includes four parts, namely data administration, data management, data alignment, and data transformation.

Data administration is the process of managing the information exchange including documentation of the source, format, context of validity, fidelity, and credibility of data. Our C2IEDM data mediation service family was originally designed and implemented with the purpose of supporting the information exchange between the multiple source database (MSDB) used in earlier prototypes, and the C2IEDM database. The insights led to a much broader concepts resulting in XML based mediation services.

In our initial example, the validity of data, which is translated from the MSDB, needs to be identified first, such as if a non-null field has null data. Next, data fidelity, i.e., whether the valid data is translated correctly, is checked based on internal relationship of each database and mapping documentation provided by our partner ACS. For example, when a non-primary data in MSDB is mapped to another non-primary field in C2IEDM, necessary primary key mapping must be issued for pursuing a non-duplicate record. Finally, data credibility, whether the data has been correctly mapped to C2IEDM, is executed, which is measured by the basic definition of C2IEDM to have the right meaningful data in the right field of C2IEDM.

Data management is planning, organizing and managing of data by defining and using rules, methods and tools. In our initial C2IEDM data mediation service, we used MSDB and C2IEDM database implementation based on MySQL. The mapping that needed to be supported by data management was twofold. First, we needed to map the Joint Command

Database (JCDB) elements plus its BML extensions used within the MSDB to the C2IEDM in order to be able to populate the C2IEDM implementation using the original BML data for validation purposes. Second, we needed to map the XML tag sets as used in the BML WSDL to the C2IEDM database in order to allow dynamic update and insert of data during runtime.

Data alignment ensures that the data to be exchanged exist in the participating system as an information entity and aggregation/disaggregating of information entities are applied so that they match the information exchange requirements, including the adjustment of data formats. In the C2IEDM data mediation service, a mapping tool called MapForce makes the whole data engineering process easier. In the evaluation phase, MapForce has been identified as our first choice for XML, database, flat file, or EDI[3] data mapping for advanced integration projects such as BML. The following figure gives a principal overview of the supported functions and formats. A description of MapForce is given in [10]. This visual data mapper can automatically generate custom mapping code in XSLT 1.0/2.0, XQuery, Java, C++, and C#. With the power to map any combination of XML, databases, flat files, and EDI to XML, databases, and/or flat files, MapForce is an appropriate tool for data management, data alignment and data transformation for information exchange.

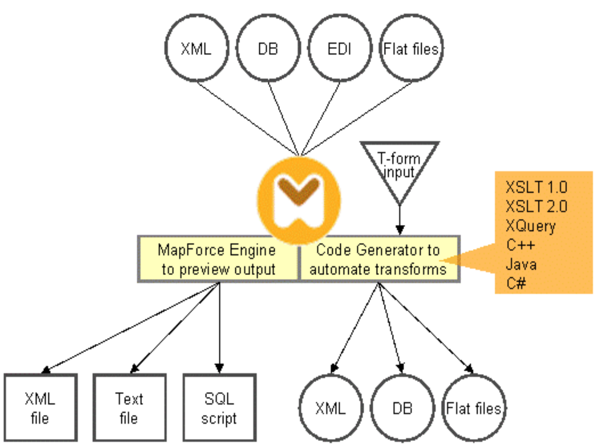

Figure 4: Functionality of MapForce [10]

[3] EDI (Electronic Data Interchange) works by providing a collection of standard message formats and element dictionary in a simple way for business to exchange data via any electronic message service.

How is MapForce used for data alignment in C2IEDM data mediation services?

First, the source and the target information description, in our initial example the MSDB and C2IEDM databases, are connected to MapForce, in our example using the MySQL ODBC driver.

Next, we use the visual mapping tool to manually draw the connections between the fields to be managed. For example, in Figure 5, the left window is showing the source database MSDB and the right window is the target database C2IEDM; the line between "eorg_index" in table "en_org_pt" of MSDB and "loc_id" in table "loc" of C2IEDM is manually drawn to identify the mapping connection.

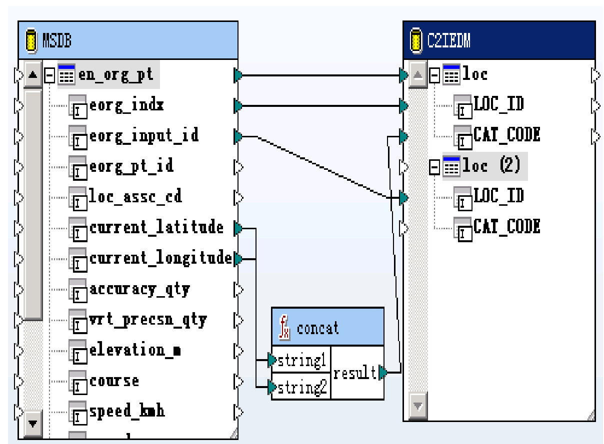

Figure 5: Example of Mapping with MapForce

In the process of data management, data are aggregated or disaggregated frequently to match the information exchange requirement when mapping from MSDB to C2IEDM. MapForce has a library, which provides functions to implement frequent data mapping needs. For example, the function "concat" used in sample 1 concatenates to strings. Other functions like "left", "right" together with "filter" and "constant" can support dividing data fields into necessary sub-components.

The following figure shows a more complex example with non-trivial functions used within the mapping. The left ten characters of field "date modified" are taken as a date, which is mapped to "PLND_START_DATE" in target database. Function "left" and "constant" are first used to take the left ten characters of "date_modified," then function "equal" and "filter" are applied to check the matched foreign key. Because table "bml2_activity" and "bml2_act" are related by foreign key "when_id," also table "bml2_activity" has a primary key mapping with table "act_task" in target database.

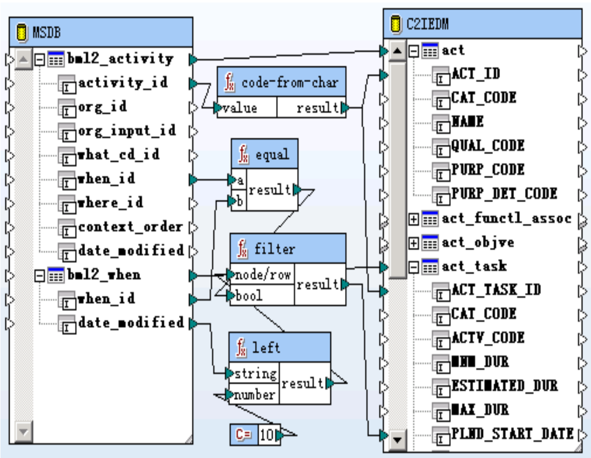

Figure 6: Example of more complex Mapping with MapForce

Removing and creating duplicated data are another usual procedure within data engineering. For many to one (m:1) mappings, which may create duplicated data, MapForce allows the user to generate duplicated tables to implement this mapping. In the third sample displayed in Figure 7, the table "loc" is produced twice to match the fields mapping from "eorg_index" and "eorg_input_id" to field "loc_id". To avoid duplicated data, whenever a non-primary field is documented to map a primary key, the necessary function needs to be set in target field. In our example, in order to avoid duplicated data from "eorg_input_id" mapped to primary key "loc_id," the table action "update if loc_id equals and insert rest" is set.

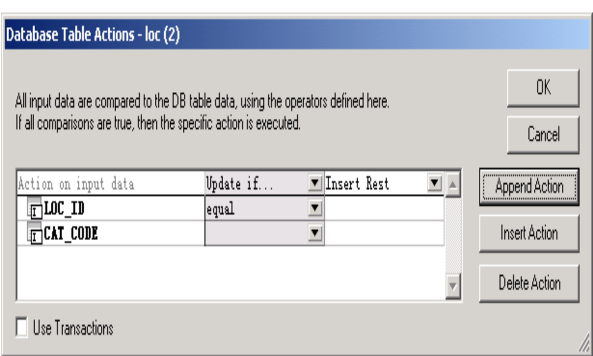

Figure 7: Setting Constraints in MapForce

Adjustment of data formats is one of most often work for data engineering. The entire primary keys from MSDB are described by string type, while integer type is applied for the primary key in C2IEDM. So, whenever there is a primary key mapping, functions like "char_from_code" and "code_from_char" are used to convert the required data type. For example, string type data from the field "activity_id" is converted through function "code_from_char" to map it to the integer type field "ACT_ID" in C2IEDM.

Data transformation is a technical process, which is usually implemented by respective algorithms. In our C2IEDM data mediation service, two methods are designed to implement data transformation.

The first method is Java based program, which is generated by MapForce based on the data engineering results described above. The resulting mapping programs are integrated into web service codes to implement the C2IEDM data mediation service.

The second method is an XSLT and Java-based mapping method set up between two XML enabled databases. For this method, XMLSpy, another Altova tool sufficient to create mature XML applications, was used to generate the source and target schemas. Using the same data engineering principles described above, XSLT codes are automatically generated by MapForce based on data management.

Although the initial examples are using two databases as source and target of a transformation, the general idea is to connect two services using different but equivalent XML dialects with each other. As already pointed out, the use of XML to describe the information exchange requirements of a service principally enables any composition of services. The use of common reference models to unambiguously define the tag sets – e.g. using the C2IEDM – ensures the semantic consistency of data exchanged. Together, these ideas can be used to be applied to instantiate mediation services as needed within service-oriented architectures using services with various data interpretations.

Generally, mediation services will navigate between individual service interpretations of data, i.e., they translate data from one interpretation into another. If a common reference model is used, mediation services can utilize the data modeling results, which map individual data interpretations to the standardized data elements of the reference model. Thus, mediation services can be applied using mediation schemas to navigate from the individual service interpretation to the standard and vice versa. This idea is captured in Figure 8.

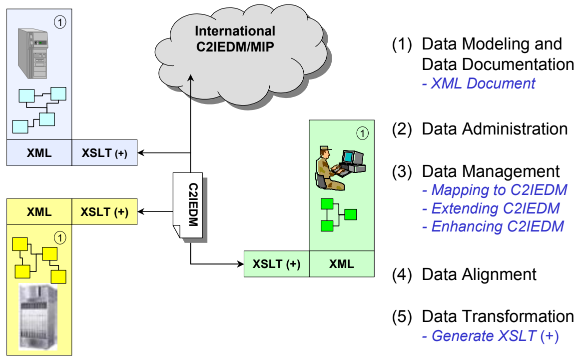

Figure 8: Conceptual View of XML based Data Mediation Service

This service family can therefore be used to migrate applications into the loosely coupled environment of services. The authors envision that such an approach may be a feasible way for the NATO Modeling & Simulation Group (NMSG) Exploratory Team ET-016 and the planned follow-on activity MSG-040. Furthermore, the approach is generally applicable to every application domain with a common reference model describing the information exchange needs as defined by the C2IEDM for military operations.

In summary, the C2IEDM storage and retrieval service family is a set of open web-based solution suggestions allowing the use of a web-based C2IEDM database with a GIG-compatible XML interface. The C2IEDM data mediation service can be used to make this database accessible for XML enabled systems which cannot yet talk C2IEDM, or it can be used for C2IEDM based information exchange independent from a database at all.

4 Current and Future Developments

We learned many lessons from and with our industry partners within the BML projects conducted over the recent months. In this section, some of these lessons learned will be described resulting in recommended current and future developments, which already are or should be conducted within research institutes to proof the feasibility of these ideas.

4.1 Graphical User Interfaces

In order to get user-appropriate access and display of the available information, the interface can easily become as important as the information itself. This calls for a major reworking and expansion of the current C2IEDM GUI.

Currently, the C2IEDM GUI only supports predefined templates that represent each of the content areas. In the next generation of the GUI, graphical rendering of the data will provide a quick reference view from which the user can extract data without viewing lengthy hierarchal listings of mission entities and objectives. The images will be able to communicate information quickly even to those that are not familiar with C2IEDM standard. It will also conform to the JIEO MIL-STD-2525 symbols as another standard that is to be employed.

The renderings can also show selected information further exploiting the previously defined content areas. The idea is to give the user only the information they want see, allowing each view of the data to be customizable to each access session. Clutter reduction also helps to make the intended purpose of the commander that much clearer, given that the content areas of the order can be reviewed individually.

Accomplishing customizability within the interface requires adding additional layers of abstraction between the data source and the final view.

- The first layer would be that of a selection interface, where the user will be able to choose the concept areas they wish to view. This selection layer will be composed of cascading services where each service will strictly comply with a single content area. It is foreseen that not all content areas will be able to be subjected to exclusion, therefore the identification of static selects (i.e. terrain), as well as dynamic selects (i.e. units, weather, etc.) will simplify this layer greatly.
- The second layer is an intermediate layer, which will identify the data selected and perform aggregation/disaggregation processes as needed. This layer preps the data, which comes from the highly normalized C2IEDM database. The layer will sit on top of the selection layer below the GUI as a means to provide last minute data checking.
- The third layer is the image preparation phase, where it will be possible to customize the display screen. It is at this phase where the information will be transformed into a format in which the graphical engine can work with.

We can utilize many ideas of our partners to develop this interface. General Dynamics/AIS only recently sponsored the development of a reconfigurable interface for heterogeneous databases. Using XML schemas, the user defines what elements he wants to display where within menus and/or popup boxes.

This software can be used to realize the ideas described here. Furthermore, the idea described as "Web Service Interest Management (WSIM)," which belongs to the main ideas developed by the XMSF partners, is published in [11], can be applied to support this development.

4.2 Complex Mapping

We do not intent to suggest that mapping is an easy process. Actually, the step data management within the data engineering process as described in [4] is very challenging.[4]

Many examples use single or basic mapping, which means that one information entity of the source is mapped to one information entity on the target side. We made the experience, and our partners shared the experience as well (see in particular [8]), that basic mapping is more the exception than the rule. What is more often observed is complex mapping, which involves records or fields that occur multiple times for a single instance, either on the target side or on the source side or on both sides. In other words: in complex mapping environments, m:n mappings are the rule. The following figure captures the early stages of a method to cope with this challenge.

- In the first step, all elements needed on the source and target sides must be identified. We assume that for the general process a top-down approach will be needed, as bottom-up approaches are not sufficient. The idea already has been coped with in [3], but a rigorous definition within a method is needed.
- In the next step, we identify the type of mapping. This mapping falls into one of two categories: 1:n mappings or n:m mappings (n:1 mappings and 1:1 mappings are trivial special cases).

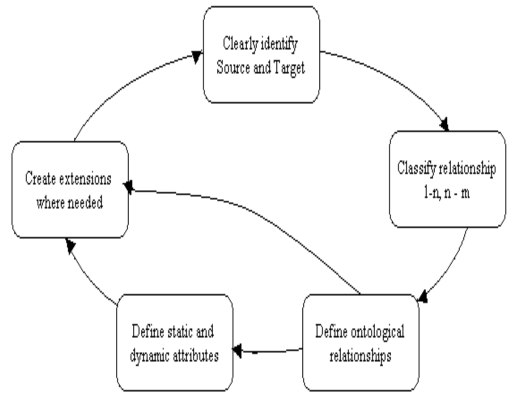

Figure 9: Complex Mapping Process

- If we will be able to define relationships not only within data models, but ultimately in ontologies (i.e., concepts explaining the application domain in a concise and unambiguous manner), this is the next step.
- The definition of static and dynamic attributes is necessary for optimization. While static attributes only must be aligned for the initialization phase, dynamic attributes change their value during runtime and are the basis for information exchange during runtime.[5]
- As pointed out in [4], the underlying ontology must be refined within the process, i.e., extensions and enhancements are necessary.

Another complete new set of challenges is introduced when information exchange is not limited to initialization tasks, but when only parts of the information are exchanged dynamically. To specify these requirements goes beyond the scope of this paper, but [3] gives some first hints by defining a sort of "business objects" which can be imported and exported by participating systems in form of propertied and associated concepts. Both families of concepts may have mandatory elements, which are not necessarily provided by the providing information exchange partner. In case of dynamic exchange, this must be solved as well. Finally, ontologies can help to define concepts of similarity to find out if a missing

[4] The interested reader is refereed to Microsoft's BizTalk Server, MSDN library, "Types of Mapping Operations", which can be accessed via this link: http://msdn.microsoft.com/library/default.asp?url=/library/en-us/sdk/htm/ebiz_prog_map_pjey.asp

[5] For an unambiguous understanding, the static parts must be communicated during the initialization. However, there is no need to exchange them during runtime as well. Example: the order-of-battle can be used for initialization, but only changes are posted during runtime.

information element may be derived from or replaced by another available information element.

This section is just a very rough sketch of very preliminary ideas; however, they show the potential and the way to go when generating a methodology, which can be applied on the big scale as well.

We see that the development of a tool-supported method that can be applied by educated engineers will be crucial to the success of BML. Without such a method, the generation of BML dialects and the necessary data engineering will be "art" solutions for individuals, but not the engineering process needed to support the armed forces.

4.3 MIP–Conformance

Without going into details, it should be stated that the C2IEDM database we currently use is semantically compliant to the MIP ideas; however, in order to become a partner within an operational environment, more constraints have to be fulfilled. The idea of data replication mechanisms (DRM) already was reflected in [2]. In general, DRM connect two or more databases and keep them consistent by updating data following protocols such as negotiated push.

ATCCIS established the ATCCIS Replication Mechanism, which is used within MIP as well. In order to participate in federations with this approach, additional metadata containing information on who has to be informed when (and many other details) has to be provided. Our database doesn't support the necessary metadata yet.

However, if the database is updated and can participate in replications based on the current ARM, every MIP system can become the source of data for BML activities. The cloud in Figure 8 could be replaced by the information sphere used by real, operational MIP systems.

It is worth mentioning that the fact that the C2IEDM is moving towards a Joint Consultation, Command and Control Information Exchange Data Model (JC3IEDM) and now even has a broader user community proofs the value of this approach. Furthermore, the NMSG ET-16 decided to build its prototype using C2IEDM compliant web services.

5 Summary and Recommendations

The two service families described in this paper are not academic toy solutions. They could be adapted to the prototype developed by our industry partners within a period of two days and were stable for the period of the demonstrations.

The authors think that the service families can be easily integrated into other systems due to their composability and scalability. We hope that many NATO partners and other interested allies will use the data mediation layer to make their XML enabled system C2IEDM compliant. The solutions leave room for improvement, but we see them as pointing towards the potential of future solutions.

However, technical solutions are not sufficient. What we need in addition to stable technical support is a set of guiding processes and principles; in other words, we need an easy to understand and easy to apply tool supported method as a basis for common engineering solutions applicable to the armed services, the joint community, and international applications. That means that the solution described here is an important step and a solution directly using data management activities to configure translation software, even in form of composable web services; nonetheless, the academic process of solving the semantic conflicts described in [4] and mapping the results remains a problem requiring domain subject matter experts to solve it.

Acknowledgements

The research work underlying this paper was partly funded by the Defense Modeling and Simulation Office (DMSO), the US Joint Forces Command (JFCOM), and General Dynamics – Advanced Information Systems (GD/AIS). The work described in this paper is part of the contribution of VMASC to the XMSF initiative conducted by the Naval Postgraduate School, George Mason University, Old Dominion University, and SAIC, and supported by sponsors in government and industry.

6 References

[1] Andreas Tolk: "Using ATCCIS as an Information Layer to couple CGF Federates and Closed Combat Simulations," Fall Simulation Interoperability Workshop, Orlando, September 1999

[2] Andreas Tolk: "Bridging the Data Gap - Recommendations for Short, Medium, and Long Term Solutions," Spring Simulation Interoperability Workshop, Orlando, March 2001

[3] Andreas Tolk: "Moving towards a Lingua Franca for M&S and C3I – Developments concerning the C2IEDM," European Simulation Interoperability Workshop, Edinburgh, Scotland, June 2004

[4] Andreas Tolk: "Common Data Administration, Data Management, and Data Alignment as a Necessary Requirement for Coupling C4ISR Systems and M&S Systems," Information & Security: An International Journal, vol. 12, no. 2 (2003): 164-174

[5] S. Carey, M. Kleiner, M. Hieb, R. Brown: "Standardizing Battle Management Language –A Vital Move towards the Army Transformation," Fall Simulation Interoperability Workshop, 2001

[6] Michael R. Hieb, J. Mark Pullen, William P. Sudnikovich, Andreas Tolk: "Developing Battle Management Language into a Web Service," Spring Simulation Interoperability Workshop, Washington, D.C., April 2004

[7] Andreas Tolk, Kevin Galvin, Michael Hieb, Lionel Khimeche: "Coalition Battle Management Language," Fall Simulation Interoperability Workshop, Orlando, Florida, September 2004

[8] Liam DeMasi, Verlynda S. Dobbs, Adam Ritchie, William P. Sudnikovich: "Implementing Battle Management Language: A Case Study Using the Command and Control Information Exchange Data Model and C4I-M&S Reference Object Model," Spring Simulation Interoperability Workshop, San Diego, CA, April 2005

[9] L. Bass, P. Clements, R. Kazman: "Software Architecture in Practice," (2nd edition), Addison-Wesley, 2003

[10] Altova Company, MapForce 2005, last accessed at Jan 2005, online available at http://www.altova.com/ products _ mapforce.html

[11] J. Pullen, R. Brunton, D. Brutzman, D. Drake, M. Hieb, K. Morse, A. Tolk: "Using Web Services to Integrate Heterogeneous Simulations in a Grid Environment," International Conference on Computational Science, Krakow, Poland, June 2004

Authors' Biographies

ANDREAS TOLK is Senior Research Scientist at the Virginia Modeling Analysis and Simulation Center (VMASC) of the Old Dominion University (ODU) of Norfolk, Virginia. He is the principal investigator. He has over 14 years of international experience in the field of Applied Military Operations Research and Modeling and Simulation of and for Command and Control Systems. In addition to his research work, he gives lectures in the Modeling and Simulation program of ODU. His domain of expertise is the integration of M&S functionality into real world applications based on open standards.

SAIKOU DIALLO is M&S Master Student at Virginia Modeling Analysis and Simulation Center (VMASC) of Old Dominion University (ODU) of Norfolk, Virginia. He implemented the web service architecture and interfaces.

KEVIN DUPIGNY is M&S Master Student at Virginia Modeling Analysis and Simulation Center (VMASC) of Old Dominion University (ODU) of Norfolk, Virginia. He implemented the C2IEDM database.

BO SUN is M&S Ph.D. Student at Virginia Modeling Analysis and Simulation Center (VMASC) of Old Dominion University (ODU) of Norfolk, Virginia. She holds a Master in Computer Science from Lamar University, Beaumont, TX. She implemented the data mediation service.

CHARLES TURNITSA is the Lab Manager for the Battle Lab at Virginia Modeling Analysis and Simulation Center (VMASC) of Old Dominion University (ODU) of Norfolk, Virginia. He managed the BML project parts done at VMASC. In addition to his work as Lab Manager, he also performs as a system engineer and researcher on several VMASC research projects. He has served in many roles as an IT professional for over a decade, performing analysis, design, and development tasks on many research projects for NASA and the US Military.