Moving towards a Lingua Franca for M&S and C3I – Developments concerning the C2IEDM

Dr. Andreas Tolk
Virginia Modeling Analysis & Simulation Center (VMASC)
Old Dominion University
Norfolk, VA 23529
atolk@odu.edu

Keywords:
C2IEDM, Data Engineering, Data Management, Data Alignment, Data Administration

ABSTRACT: *Under the umbrella of NATO, the ATCCIS Permanent Working Group developed the Army/Allied Tactical Command and Control Information System (ATCCIS) Data Model for a couple of years, establishing an unprecedented amount of international agreement on information exchange requirements on the battlefield. The model was recently established as the Allied Data Publication Standard No. 32 (AdatP-32) under the name "Land Command and Control Information Exchange Data Model (LC2IEDM)." As the concepts are not bound to land services, the name was changed to "Command and Control Information Exchange Data Model (C2IEDM)." An alternative name used referring to the same model is "Generic Hub" or "Battlefield Generic Hub." Due to its high level of maturity, this data model also became the data model of the Multilateral Interoperability Program (MIP), established by NATO to interconnect their C3I systems.*

Although various prototypes using the ATCCIS/C2IEDM data model to connect C3I and M&S systems have been presented to SISO in the recent past, the concept is approaching a new quality, as various projects and concepts are starting to adapt the ideas. Following some general concepts dealing with the challenges of correct linguistic information mapping, this paper will give a short history of the development and first prototypes and an overview on actual projects, in particular the Extensible Battle Management Language (XBML) project. Furthermore, it will extrapolate how – if the proposed route is followed – the ATCCIS/C2IEDM data model can become the backbone for a common ontology for military information exchange in coalitions of services, nations, and system families, such as C3I and M&S.

1. Introduction

The "Command and Control Information Exchange Data Model (C2IEDM)" may best be described as an idea whose time has finally come.

The idea itself is trivial: In order to have an unambiguous specification of information exchange requirements (IER), this specification must be standardized. However, before the C2IEDM could begin its march towards victory, many prejudices had to be overcome.

The challenge with IER is that they change as rapidly as do the application domains. To standardize IER in a dynamic environment requires a lot of modeling expertise on the one hand and a lot of expertise in the application domain on the other. Many experts therefore expected this to be a never-ending task and preferred to implement working one-to-one solutions.

Furthermore, many system developers refuse to evaluate IER standards because they misunderstand that the IER standards don't mandate the implementation of their application. Agreeing to use a given IER protocol doesn't mean that the objects to be exchanged must also be used internally.

The NATO environment delivered the necessary ingredients for a successful approach to cope with this issue: short notice collaborations using multinational and multilateral connections, often limited to legacy systems, delimited by decreasing budgets for continually growing numbers of interfaces in the operational contexts of multi-service, multi-nation, multi-echelon, and ill-defined task constraints all over the world. Military heterogeneity could not be better defined.

Not only became clear that soldiers would need to operate with their own systems in such an environment, it was also obvious that the introduction of one common system to fulfill everybody's needs was not technically feasible.

Therefore, a common approach merging the existing systems and building a solid basis for future developments, in the sense of additional integrations or common new procurements to improve the existent architecture, was needed. To this end, the approach to use a common information exchange model for all participating command and control systems was defined. Over the last couple of years, the data model improved and evolved into the current NATO Allied Data Publication Nr. 32, the *Command and Control Information Exchange Data Model* (C2IEDM).

Today, NATO utilizes this data model successfully in two application domains, namely data management and information exchange. Within the United States, the interest in this matured approach is steadily increasing. Applications are no longer limited to prototypes or academic feasibility studies, but operational systems are managed and coupled using the approach discussed here. To show how these efforts can be particularly used to couple M&S and C4I systems, this paper will discuss

- the main issues when using a reference data model to align data models
- the history and anatomy of the C2IEDM
- the use of the C2IEDM in the context of the extensible Battle Management Language (XBML) project
- migration rules for legacy M&S and C2 systems to use the C2IEDM for information exchange.

This paper can neither replace the detailed study of the C2IEDM references nor can it explain the Battle Management Language and its application. However, there will be enough sources given to enable the interested reader to deepen his knowledge in these domains. This paper shall give an overview on the management and decision level helping to decide whether the ideas are suitable for a given domain or project or not.

2. Data Engineering

The following definitions are taken from [1] in which recent works concerning data management and data alignment dealing with C2 and M&S interoperability were summarized.

2.1 Components of Data Engineering

Common to all current solutions dealing with C2 and M&S interoperability is that the system designer tasked with the integration has to know *what* data is located *where*, the *meaning* of data and its *context*, and into what *format* the data have to be transformed to be used in respective distributed applications within the overall system. The objective of data administration, data management, data alignment, and data transformation is to generate the answers to these questions. They can be defined as the building blocks of a new role in the interoperability process: the tasks of *data engineering*. Of these tasks, the first three can be standardized and used in a general manner. Only the task of transforming the data is really system dependent. It is interesting that most efforts presented in the recent Simulation Interoperability Workshops fall into this category while the essential tasks of data administration, management, and alignment are considered to "just happen somewhere else effectively" and are of no concern to the interface developer. However, the author is convinced that theses steps of data administration, management, and alignment are necessary first achievements in an overarching interoperability framework as envisioned in [2].

In the context of [1], the terms are defined as follows:

- *Data Administration* is the process of managing the information exchange needs that exist within a group of systems, including the documentation of the source, the format, context of validity, and fidelity and credibility of the data. Data Administration therefore is part of the overall information management process.
- *Data Management* is planning, organizing and managing of data by defining and using rules, methods, tools and respective resources to identify, clarify, define and standardize the meaning of data as of their relations.
- *Data Alignment* ensures that the data to be exchanged exist in the participating systems as an information entity or that the necessary information can be derived from the data available, e.g., using the means of aggregation or disaggregation.
- *Data Transformation* is the technical process – often implemented by respective algorithms within the gateways and interfaces – of aggregation and/or disaggregation of the information entities of the embedding systems to match the information exchange requirements including the adjustment of the data formats as needed.

Many efforts are focusing on data transformation, i.e., the programming or maintenance of interfaces. However, if such efforts are not accompanied by an alignment of the respective management processes for data administration, management, and alignment, the result is in the best case a temporary valid solution that is effective until the next update of one of the participating systems. Consequently, the managing processes of the participating systems must at least be harmonized. In the ideal case, the program managers will even use the same methods and supporting tools to do so under a common, overarching approach.

The C2 community as well as the M&S community is currently in the process of establishing solutions supporting these management efforts. In order to insure continuous interoperability, these processes must be harmonized and coordinated, leading to a common approach which is referred to as common data engineering. How this can be done, is generally described in [2] and technically specified in [3] as follows (see also section 2.3):

- Every component delivering operationally required functionality defines its input and output parameter in form of XML descriptions of the interface. XML becomes the common syntax to describe data to be exchanged in a common technical language.
- Data administration can be directly supported by web services when every data source comes with a web service specifying source, format, context of validity, and fidelity and credibility of the data. These data administration information web services have to be published using a data administration specified universal description, discovery, and integration (UDDI) registry.
- As data management will be dealt with in the following sub-section, here is the main idea: As the standard reference model comprising the standardized data elements (SDE) as well as the data model to be managed both are described in the XML standard, mapping of the models equals mapping of the tag sets used.
- After the models have been mapped to the standard reference model, data alignment is a matter of one-on-one comparison and can be done automatically.
- Furthermore, as mapping equals tag set mapping, data transformation can be supported by Extensible Stylesheet Language Transformations (XSLT) applications.

To summarize, data management remains as the only academic process in the data engineering chain. All other steps can be automated based on the agreement to the right standards. The following section shows how a reference model such as the C2IEDM can support data management.

2.2 Model-based Data Management

As stated before, data management is planning, organizing and managing of data by defining and using rules, methods, tools and respective resources to identify, clarify, define and standardize the meaning of data as of their relations. This can be done by point-to-point mappings or by using a standard reference model.

There are many challenges that have to be dealt with when mapping models to each other. In the model-based data management approach, we are dealing with mapping concepts to each other. The following four classes are defined in [4]:

- *Semantic Conflicts:* the concepts of the different schemata do not match exactly, but have to be aggregated or disaggregated. They may only overlap or be subsets of each other, etc.
- *Descriptive Conflicts:* there are homonyms, synonyms, different names for the same concept, different attributes or slot values for the same concept, etc.
- *Heterogeneous Conflicts:* the methodologies being used to describe the concepts differ substantially, e.g., one concept is described in the Unified Modeling Language (UML), the other in the relational data model description methodology IDEF1X.
- *Structural Conflicts:* different structures are used to describe the same concept, e.g., in one local schema an attribute is used, in the other schema a reference to another concept is used to describe the same part of the view of "reality".

When using a standard reference model in a standard description schema based on XML, the issues of descriptive and heterogeneous conflicts are solved per definition. Structural conflicts can be solved by introducing the ideas of properties, which are characterizing descriptions or capabilities, and propertied concepts. The same properties can be described in various forms; however, these different structures can be mapped to the same properties and propertied concepts in the process of the first step of data management. When consistent design patterns are applied, this solves the issue of structural conflicts. (Remark: The Platform Independent Models (PIM) used within the Model Driven Architecture (MDA) are following the same idea; see [2] for details.)

What remains is the category of semantic conflicts. Although in the business domain first efforts are conducted to support automatic translation of XML schemas into each other [5,6], the author is convinced that the domains of C2 and military M&S are too complex to allow automatic translation without a human in the loop. However, the idea to have a conceptual model as a basis for mapping, as recently published in [6], supports the model-based data management approach recommended in this paper. In order to describe possible semantic conflicts, we use the terminology introduced in [7]:

- *Associated concepts* are semantic entities in which data is given in a context. Within data models, these are the replication domain sets, which are those tables connected by relations that have to be sent from one instance to another in order to make sense for the user. In the XML world, this can be mapped to XML documents satisfying the XML schema. In the relational database world, these are tables and associations fulfilling the requirements for referential integrity. In some newer publications, these ideas are alternatively referred to as domain-specific ontology layers above the propertied concepts, as these constructs are needed to provide the domain-specific context of the data.

- *Propertied concepts* are a collection of specifying characteristics for an entity in the domain of knowledge. In ontologies using data models to structure its information, this can be mapped to table and its attributes. Within XML, this is the collection of XML tag sets. Within relational databases, these are tables defined by their attributes.
- *Property values* are the allowed values for a specifying characteristic. Of particular interest are enumerations. Within XML, these are the allowed values within the documents. Within relational databases, these are enumeration values for attributes.

When applying these terms, the following main semantic conflict categories can be described:

- *General aggregation problems*: if two data models to be mapped are on totally inadequate resolution levels, associated concepts of the high-resolution model are likely to be mapped to property values on the highly aggregated model. For example, the high-resolution model may describe weapon systems as associated concepts of platforms, weapons, and sensors (all being propertied concepts) while the highly aggregated model only counts the number of weapon systems within a simulated unit.
- *General alignment problems*: not all information necessary in the target model is supplied by the source model. This is a problem of insufficient data modeling and not of data management or data alignment. This problem has to be solved by the system developers. If the missing data really is important, the system has to be changed. Data management can only show the problem, but cannot create a solution.

What can be dealt with effectively and efficiently by data management are the aggregation problems dealing with composing and decomposing information elements, which can be subdivided further into the following cases:

- *Extension of property values*: the property values of the model to be mapped exceed the property values of the reference model. In order to cope with the additional value, most likely additional enumerations, the reference model must be extended to comprise these additional property values in respective properties. In some cases, the introduction of additional properties may be necessary (see extension of propertied concepts).
- *Enhancement/refinement of property values*: the resolution of the model to be mapped is higher than the resolution of the reference model. In this case, the resolution of the reference model has to be increased in order to be able to cope with these higher detailed values. In addition, the mapping functions of other mapped models have to be adapted to the new situation (aggregation and disaggregation to the new resolution of the property values).
- *Different grouping of property values*: although the property values are the same they are used to describe different propertied concepts. This is a true semantic conflict, as the “specifying characteristics” are specifying different concepts in both models. Normally, this conflict can be resolved by enhancement/refinement of the affected propertied concepts.
- *Extension of propertied concepts*: a propertied concept of the model to be mapped comprises additional properties. If this doesn’t lead to an enhancement/enrichment, the number of properties has to be increased to cope with these information elements.
- *Enhancement/refinement of propertied concepts*: the resolution of the model to be mapped is higher than the resolution of the reference model. In this case, the resolution of the reference model has to be increased in order to be able to cope with these higher detailed concepts. In addition, the mapping functions of other mapped models have to be adapted to the new situation (aggregation and disaggregation to the new resolution of the propertied concept).
- *Different grouping of propertied concepts*: although the propertied concepts are the same they are used to describe different associated concepts, which means, they describe different semantic concepts, such as fields necessary to describe an action like an attack, etc. This is a semantic conflict on a higher level and very common, as the associated concepts – or ontology layers – are describing the semantic concepts of the underlying model or application domain. What is seen as mandatory in the context of one model can be seen as irrelevant in another model. There are two solutions: (1) The reference model is used as the standard to be enforced for all participating models, which can lead to mentionable re-implementations on the side of the participating models; or (2) modeling associated concepts, but not using them to mandate semantic concepts to be accepted by all models.
- *Extension of associated concepts*: if the reference model has more propertied concepts in the associated concept, the solutions described under

different grouping of propertied concepts have to be applied respectively: if associated concepts are mandated, the models to be mapped have to be enhanced. If the model to be mapped has more propertied concepts, it must be decided if these are model specific issues or if the semantic concepts of the reference model have to be extended. If they are extended, all former models must be checked to see if they fulfill these extended semantic requirements.

- *Enhancement/refinement of associated concepts*: these conflicts can normally be solved by increasing the resolution of the reference model by splitting the propertied concepts into new, higher resolution propertied concepts. In addition, the mapping functions of other mapped models have to be adapted to the new situation.

If (a) propertied concepts are mapped to property values of the reference model, or (b) associated concepts are mapped to property values of the reference model, or (c) associated concepts are mapped to propertied concepts of the reference model, the resulting challenges can be coped with using the methods of grouping, extending and enhancing described above. If in the three mapping challenges the model to be mapped has the higher resolution resulting in such challenges, this may be a hint that the reference model is not applicable for mapping.

Finally, it should be pointed out that the standard reference model is used to define the information exchange between the components and does not imply its use as a common data model for all components. Actually, there is no need for a physical implementation of this model outside the data management activities. After two models have been mapped to the reference data model, the results can be used to generate an XSLT schema connecting the two systems directly.

2.3 Data Engineering and XML

As previously pointed out, the use of XML as the common syntax or the container used to describe the models to be mapped to each other is recommended in [2] for military M&S components. A more general view on this issue is given in [8], and some of the main arguments shall be given here as well:

- XML is a strong candidate to mediate structured and semi-structured data. XML supports advanced views, intelligent agents, and generally heterogeneous application integration.
- XML data structures support data mediation because XML tags are semantic free, XML completely separates data from presentation, and XML supports internationalization and media independence.
- XML promotes universal data access from alternate user interface metaphors, which is of particular interest for data administration.

Of additional particular interest is the potential to combine software agents and XML interfaces to data. In combination with a reference data model to define the semantics of the tags, agents can use XML to get access to heterogeneous data, can map and combine them, and become an intelligent access layer. XML in combination with intelligent agents are key technologies for the semantic web.

After the general challenges have been defined in this section, the C2IEDM and its use for model based data management in general – and in particular for M&S and C2 system interoperability – will be dealt with.

3. Overview of the C2IEDM

Within this section, the use of the C2IEDM for data management and data management and information exchange will be presented. After presenting a short overview on the history and the anatomy of C2IEDM, the theoretic aspects of the last sections will be reflected describing the use of C2IEDM within these two application domains.

3.1 History of the C2IEDM

In 1978, NATO's Long-Term Defense Plan (LTDP) Task Force on Command and Control (C2) recommended that an analysis be undertaken to determine if the future tactical Automatic Data Processing (ADP) requirements of the Nations, including that of interoperability, could be obtained at a significantly reduced cost when compared with the approach that had been adopted in the past. In early 1980, the then Deputy Supreme Allied Commander Europe initiated a study to investigate the possibilities of implementing the Task Force's recommendations. This was the birthday of the ATCCIS Permanent Working Group (APWG), which was dealing with the challenge of the future C4I systems of NATO.

The ATCCIS approach comprised more than just another data model. It was designed to be an overall concept for the future C4I systems of the participating nations. One of the most important topics of ATCCIS was that each nation could still build independent systems with their own "view of the world" and respective applications, business rules, implementation details, etc. Thus, ATCCIS defined a common kernel to facilitate common understanding of shared information and, therefore, facilitated facing the general challenge to reach interoperability based on various heterogeneous IT solutions. The Army/Allied Tactical Command and Control Information System (ATCCIS) comprised

- the ATCCIS data model (including a standardized common generic hub and sub-functional areas of national concern),
- the ATCCIS system architecture (with a kernel of common access points to the logical ATCCIS data model on the one side, and access points to standard communication protocols like TCP/IP on the other),
- the ATCCIS Information Resource Dictionary System (AIRDS) with references about information and information structure and context for each data element, and
- the ATCCIS Replication Mechanism (ARM) allowing internal communication by user driven and specified database replication between two ATCCIS compliant systems.

The technical feasibility was demonstrated several times and ATCCIS based systems were a reliable part of the annual Joint Warrior Interoperability Demonstrator (JWID) programs. Finally, the ATCCIS data model became a NATO standard with the Allied Data Publication Nr. 32 (ADatP-32) with the new name Land Command and Control Information Exchange Data Model [9], which was adapted in 1999.

In parallel to this, the Multilateral Interoperability Program (MIP) was established by the project managers of the Army Command and Control Information Systems (C2IS) of Canada, France, Germany, Italy, the United Kingdom and the United States of America in April 1998 in Calgary, Canada. MIP replaced and enhanced two previous programs: BIP (Battlefield Interoperability Program) and QIP (Quadrilateral Interoperability Program). The aim of these programs was similar to the present MIP but each was active at a different level of command.

By 2002, the activities of ATCCIS/LC2IEDM and MIP were very close, expertise was shared, and specifications and technology were almost common. The merger of ATCCIS and MIP was a natural and positive step and this was recognized by the almost immediate publication of a NATO policy that endorsed MIP. LC2IEDM became the data model of MIP, which established the Message Exchange Mechanism (MEM) and the Data Exchange Mechanism (DEM) based on replication mechanism.

Finally, in 2003 the name was changed to Command and Control Information Exchange Data Model (C2IEDM). The most recent version is the Generic Hub version 6.1 and it can be downloaded from the MIP website [10]. This site also comprises a lot of additional information on the history and use of the recent C2IEDM developments.

3.2 Anatomy of the C2IEDM

The C2IEDM data model comprises two categories: the Generic Hub (GH) and the Sub-Functional Areas (SFA).

The underlying idea is that a data model driven by C2 user requirements must encompass information from multiple functional areas in the domain of military operations. Consequently, a C2 data model serves as a "hub" for unifying information concepts that are embodied in the data specifications of functional areas. The desired goal in the long term would be a federation of data specifications using the C2 data model as the basis for functional area models. This would ensure that the data that is common between the functional areas, the hub, is viewed and structured in a standard way, and that the data model views can be readily integrated into coherent structures wherever such integration is needed. All common data, or better said all data that need to be exchanged by at least two functional areas, become part of the Generic Hub (GH). The remaining data is modeled as a specific extension of the Generic Hub data into the Sub-Functional Areas (SFA).

Initial evolution of the C2IEDM under MIP included specific inputs from the following functional areas: conventional fire support, barrier engineering operations, communications and electronics, and personnel administration. Operational requirements have been drawn from these as well as other areas, as documented on [10].

To summarize, the C2IEDM GH is intended to represent the core of the data identified for exchange across multiple functional areas. It lays down a common approach to describing the information to be exchanged in the command and control domain. Thus, the approach is generic and not limited to a special level of command, force category, etc. In general, C2IEDM describes all objects of interest on the battlefield, e.g., organizations, persons, equipment, facilities, geographic features, weather phenomena, and military control measures such as boundaries, using a common and extensible data modeling approach.

The ATCCIS Generic Hub is based on information concepts, also referred to as the 15 independent entities. Five key information concepts are of fundamental importance in generating the structure of the data model. They are defined in Table 1.

The distinction of objects and items is essential. The battlefield consists of a large number of objects, each with its own set of characteristics. Objects may be described as a class or type rather than as individually identified items. Actual instances are catered for by use of OBJECT-ITEM. Types are recorded as OBJECT-TYPE. While general attributes are collected on the type side, such as general capabilities and abilities, only the instantiation specific values are on the item side. Examples are the caliber of the weapon being specified on the type side, but the actual ammunition state and location are on the item side.

Another important issue enabling the C2IEDM GH to be extended as required is an extensive use of the categorization mechanism. To show the principle, the following Table 2 shows the category codes applicable to OBJECT_ITEM. In recent versions, the category code UNKNOWN was introduced for an OBJECT-ITEM, which is tracked but has not yet been classified. If new information has to be introduced, this can be done by adding new attribute values (in particular new category and sub-category codes), adding new attributes, adding new tables, or adding new associations (compare to section 2.2)

Concept	Definition
OBJECT-ITEM	An individually identified object that has military significance. Examples are a specific person, a specific item of materiel, a specific geographic feature, a specific coordination measure, or a specific unit.
OBJECT-TYPE	An individually identified class of objects that has military significance. Examples are a type of person (e.g., by rank), a type of materiel (e.g., self-propelled howitzer), a type of facility (e.g., airfield), a type of feature (e.g., restricted fire area), or a type of organization (e.g., armored division).
CAPABILITY	The potential ability to do work, perform a function or mission, achieve an objective, or provide a service.
LOCATION	A specification of position and geometry with respect to a specified horizontal frame of reference and a vertical distance measured from a specified datum. Examples are point, sequence of points, polygonal line, circle, rectangle, ellipse, fan area, polygonal area, sphere, block of space, and cone. LOCATION specifies both location and dimensionality.
ACTION	An activity, or the occurrence of an activity, that may utilize resources and may be focused against an objective. Examples are operation order, operation plan, movement order, movement plan, fire order, fire plan, fire mission, close air support mission, logistics request, event (e.g., incoming unknown aircraft), or incident (e.g., enemy attack).

Table 1: The ATCCIS Generic Hub concepts

Every concept is further defined by an unambiguous set of properties. To implement the categorization mechanism just introduced, C2IEDM uses category codes and sub-category codes utilizing well-defined enumerations. The meaning of each numeration – often including the source of the definition – is specified as well in the documentation.

Finally, the concept of replication domain sets must be explained. The concepts of C2IEDM are connected by relations. Some of these relations are mandatory in order to assure that within a semantic association of information spread over various tables all pieces of data are presented. Every one of these semantic concepts forms a replication domain set; in other words, the set is defined as a group of associated concepts connected by mandatory relations. The concepts comprised are perceived to be fundamental to command and control and must be provided by every participating system. For example, every message has an originator and a time stamp, so every replication set must comprise this information as well.

Concept	Definition
FACILITY	An OBJECT-ITEM that is built, installed, or established to serve some particular purpose and is identified by the service it provides rather than by its content (e.g., a refueling point, a field hospital, a command post).
FEATURE	An OBJECT-ITEM that encompasses meteorological, geographic, and control features that are associated with a location to which military significance is attached (e.g., a forest, an area of rain, a river, an area of responsibility).
MATERIEL	An OBJECT-ITEM necessary to equip, maintain, and support military activities without distinction as to its application for administrative or combat purposes (e.g., ships, tanks, self-propelled weapons, aircraft, etc., and related spares, repair parts, and support equipment, but excluding real property, installations, and utilities).
ORGANIZATION	An OBJECT-ITEM that is an administrative or functional structure.
PERSON	An OBJECT-ITEM that is a human being to whom military significance is attached.

Table 2: The OBJECT-ITEM categories

3.3 Applications of the C2IEDM

Without going into detail, there are two application domains for the C2IEDM: data management and information exchange.[1] Both domains are coming together under MIP within NATO. Furthermore, national applications are contributing as well, e.g., the German concept of a future M&S Network is based on a methodology of knowledge and information management. The Integrated Army M&S Data Network will be established as an operational and technical concept to allow knowledge transfer between operational headquarters, study facilities, procurement agencies and training installations. The C2IEDM is the core data model within these efforts [11].

In order to manage the various information exchange requirements between the heterogeneous solutions of the participating nations, NATO had to establish a

[1] The two domains influence each other. In [2] a way is shown to configure XML based mediation layers as suggested in [5, 6, 8] applicable in the military domain.

flexible and efficient NATO Data Administration Group (NDAG). It was their task to plan, organize and manage the information exchange requirements by defining and using rules, methods, tools and respective resources to identify, clarify, define and standardize the meaning of data as of their relations. The NDAG did this by using the C2IEDM as the common reference model. Each participating data model was mapped to the properties, propertied concepts, and associated concepts of the C2IEDM in order to generate standardized data elements, which were unambiguously defined and it was mandatory to use them for data exchange. When the ATCCIS/LC2IEDM projects were merged, the NDAG activities became part of the MIP group. However, the results are valid beyond the C2 systems participating in MIP. They are also used to define new systems and systems procured by NATO whenever applicable.

Today, the second domain is tightly connected to MIP. Over the last 10 years, ATCCIS demonstrated the technical feasibility in several prototypes and demonstrations. With MIP, however, a new level of maturity is reached. A detailed overview on supported and desired information exchange requirements, including guidelines and definitions for the MIP Tactical C2IS Interoperability Requirements (MTIR), can be downloaded from [10].

In September 2003, MIP conducted an Integrated Operational Test and Evaluation Exercise at the C2 Support Center in Ede, The Netherlands. The intention of this event was to prove and demonstrate interoperability of information between the National C2 systems of 12 members. This capability gap is growing in significance for NATO. The significance lies in the nature of the capability delivered, which is that C2 systems that are otherwise completely different from each other may, nevertheless, interoperate and share information using the C2IEDM for information exchange.

Although further development is required in certain areas, such as operational procedures for information exchange over the MIP interface need to be further defined, agreed upon and incorporated into national SOPs, and the MIP technical specifications need to be "tightened up," in general the MIP interface performed well and supported the exchange of data necessary to establish the Common Operational Picture. However, as national C2 systems differ in capability they were not able to utilize all exchanged data. The detailed report on IOT&E was prepared with a meticulous analysis of the data captured during the scenario.

Within the United States, the Department of Defense is setting up an XML repository with recommended tag sets to be used [12]. The Institute for Defense Analysis and the Naval Postgraduate School developed an XML tag set based on the C2IEDM Generic Hub Version 5.0, which was submitted as a recommendation for coalition operations to the repository. The XMSF projects described in the following section used the same methods and techniques to generate the necessary XML tag set based on the actual version C2IEDM Generic Hub 6.1.

In summary, the C2IEDM has been proven to be a valuable contribution to interoperable solutions within NATO and the application domains of data management and information exchange have been proven to provide valuable methods supporting general concepts facilitating collaboration based on heterogeneous solutions.

4. C2IEDM within XBML

This section is not a summary of descriptions of the Extensible Battle Management Language (XBML) project. The necessary descriptions can be found in papers explicitly dealing with this topic, for example in [7]. In the scope of this paper, the use of the C2IEDM within XBML and what migration rules must be applied in order to become part of XBML will be described.

4.1 Multi Source Data Bases in XBML

The Extensible Battle Management Language (XBML) program combines the technical flexibility of the Extensible Modeling & Simulation Framework (XMSF) with the innovation of the unambiguous Battle Management Language (BML). The concept important for this paper is that C2 components, such as the Combined Arms Planning and Execution System (CAPES), and simulation components, such as the One Semi-Automated Forces Test Bed (OneSAF TB) system, communicate via XML messages using a common Multi-Source Database (MSDB). The BML component checks the consistency, and enhances and extends the orders. To summarize it, BML blends the structure that allows automation of the language and ease of use for the military professional. BML is an evolution in the command language that provides a means to gain structure while remaining transparent to the military user. It is based on doctrine and linked to the doctrinal sources, both to ensure standard use/understanding, and to foster concise and precise use of the language. The output of the automated system is dependent on whether the intended audience is a human, a software "intelligent agent" or an autonomous robot. However, the component of main importance in this paper is the MSDB.

The MSDB must support the structure of the language used to exchange the information between the components. Actually, the MSDB must support all languages of the participating components, which means that data engineering – in particular data management and data alignment – is necessary. While the first U.S. Army sponsored prototype of BML utilized the Joint Common Database (JCDB) to implement the MSDB, within the XBML project the migration towards the C2IEDM seemed to be the natural choice, as XBML targets the joint and combined community. In other words, XBML shall become an integration platform for C2 and M&S components of all services and all allies.[2]

The technical issue of data administration is coped with the requirement that the data to be exchanged, i.e., data to be used to populate the MSDB or to be extracted from the MSDB, must be accessible in the form of XML files. After data management, which is done by mapping the models to the C2IEDM, data alignment is a matter of checking if all target sets are obtainable on the source site. As target and source are XML structures, data translation is not an issue.

Data management is the real academic and semantic challenge. Even when data is aligned, every component is likely to have a slightly different interpretation. While orders without an originator, receiver, or a timestamp are not perceived to be valid on the C2 side, this is often the case on the M&S side. The question to be solved is, which side should be the standard for the mapping. If the side with the higher detail and the more constraints sets the standard, the other side must at least be enhanced, maybe even re-implemented. If the less constraint system set the standard, the result may not be accepted on the other side. Using the example of C2IEDM this means that M&S systems – or non-compliant C2 systems as well – must be adapted to support all constraints of the model. If not all constraints will be supported this will result in a population of the C2IEDM that can not be the source of a DEM or MEM as defined in [10]. However, as long as full C2IEDM compliant systems are only used as sources and not as targets, the challenge of fulfilling all constraints is not an issue.

These evaluations lead to a set of gradually increasingly demanding migration rules for components to be integrated into the XBML framework as defined in the next subsection. The level of integration is driven by the necessary degree of interoperability between the components to be integrated.

4.2 Migration Rules

The migration to the C2IEDM compliance within XBML is conducted in phases; the gradually increasingly detailed migration rules can be applied to integrate future components as well.

- *Level 1: Common Syntax based Coupling*. The loosest way to couple components is by simply exchanging information using the same medium and syntax. XBML requires XML structured information. The components agree on an XML structure to use to exchange data. The mapping to this data is within the responsibilities of the system developers.[3]
- *Level 2: Common Dictionary based Coupling*. The data elements to be exchanged are unambiguously defined by pointing to a definition within a common dictionary, which is gradually enhanced and extended by the mapping processes.

  The first BML prototype used the vocabulary of the Army Field Manual (FM 101-5-1, future FM 1-02), which was mapped to the data elements of the JCDB (but it required 113 additional tables to cope with all information). XBML uses the C2IEDM definitions. When applicable, new entries will use already agreed definitions for the new data elements. [4]

  In practice this means that the same XML structures as used on level 1 can be still used, but the tag set names are references to well-defined elements of the dictionary, or in other words, they are names of properties within propertied concepts of the reference data model.
- *Level 3: Common Low Level Semantics.* While on level 2 the lexical meaning of the properties and propertied values are used to tag the information to be transmitted, the level three introduces semantic meaningful concepts. A propertied concept groups properties together, which describe the same concept, i.e., something which has a meaning in the simulated or real world. Similarly, associated concepts group concepts together, or build domain-specific ontology layers. On level 3, this semantic structure of the reference model is used to bind the properties of the target

2 From the technical standpoint, even broader inter-agency or homeland security applications are possible. However, the integration may create semantic challenges not too easy to handle.

3 This is very similar to the development of distributed simulation systems based on the High Level Architecture (HLA), where within the federation development a Federation Object Model (FOM) has to be agreed on, and to which the simulation system developer must map the data of the simulation.

4 The vocabulary comprises terms from the Field Manual of the U.S. Army (FM 101-5-1, future FM 1-02) and the Joint Publication 1-02, Department of Defense Dictionary of Military and Associated Terms. The future vocabulary will include Service specific dictionaries expanding the syntax and semantics as defined through the UJTL, CJCSM 3500.04B, and Joint Doctrinal manuals, as well as each Service's task list (Universal Naval Task List (UNTL), OPNAVINST 3500.38/MCO 3500.26/USCG COMDTINST M3500.1, Air Force Task List (AFTL), Air Force Doctrine Document 1-1) and their respective doctrinal manuals.

and source databases. In practice this means that an implementation of the reference model can be used to store data to and retrieve data from; in other words, an extension of the C2IEDM can be used to implement the MSDB.

- *Level 4: Common High Level Semantics.* While on level 3 the properties, propertied concepts, and associated concepts of the reference model were used to model the information, on level 4, they are mandated. In other words, the referential integrity of mandated values and mandated associations in the reference model is binding for the target and source model. The target model must be able to fulfill all constraints when populating the MSDB, and the source database must be able to obtain all data from such a data source. In practice, the MSCB based on the C2IEDM becomes a target and a source for real C2 components using the DEM or the MEM as defined within MIP.

Within all these efforts, the guiding principle must be to reuse as much of the already incorporated information as possible. The objective of model-based data management is the definition of unambiguous data elements in the form of property values, properties, propertied concepts, and associated concepts. Creating new concepts comprising data elements that already exist somewhere else is against this very basic rule, because it creates ambiguity. This is particularly important in level 3 and level 4 applications.[5]

What does this mean for migration rules? Components to be integrated into this environment, such as the components of the recently presented France system APLET [13], must follow the following steps:

1. Information Exchange Requests must be described in the form of XML structures.
2. Information Exchange Requests and describing XML structures must refer to definitions within a gradually increasing XBML dictionary.

   If the necessary term is not yet in the dictionary, a domain standard to define the term should be used, such as the Joint-Pub 1-02, and the dictionary must be extended.
3. The terms of the XBML dictionary are structured in form of the C2IEDM – extended and enhanced by data management activities. The XML tag set must point to the definitions within the C2IEDM structure.
4. The information contained in one information exchange requirement, such as an attack order, which is normally contained in one XML structure as it is a semantic logical unit of the data model, is modeled using the properties, propertied concepts, and associated concepts of the C2IEDM. It may be necessary to fill in some additional values in order to associate concepts comprising data from the target that are not considered to belong to the same semantic logical unit in C2IEDM.
5. Finally, the information exchange request has to fulfill all constraints of the C2IEDM, which means that the information exchange request must be replication domain sets.

Which level of C2IEDM compliance is appropriate for a project depends on the intended use: only if a fully C2IEDM compliant system is the target, the highest level of compliance is necessary.

Finally, it is worth mentioning that all these efforts can at best ensure interoperability up to the semantic level. It will remain necessary to continue and work on the issues of dynamic/pragmatic interoperability and finally conceptual interoperability in follow-on activities [14].

5. Summary

The C2IEDM is a matured data model suitable to couple C2 systems. Whenever it is essential to "speak to C2 systems in their own language" – that means to fulfill their data needs and meet all their constraints and assumptions, such as that data are normally connected with the ideas of reports or messages and therefore have sources, timestamps, etc. – the use of the C2IEDM with all referential integrity mandates makes sense. Whenever the semantic consistency is the main issue, and the strict C2 implied referential integrity must not necessarily be assured – for example, when information is only extracted from a C2 system to feed an M&S system – the C2IEDM can be used as an outstanding "dictionary" unambiguously defining the terms used to describe military operations in all domains. The combination of the XML standard with the C2IEDM as namespace and repository source is a powerful tool ensuring interoperability for legacy and future systems and applicable to C2 and M&S system as well.

Finally, it should be pointed out that C2IEDM does not mandate the use of a common data model; its focus is much more a common information exchange model in the sense of service interface definitions. The fact that

[5] We have to distinguish between "views", which must be application specific, and the place of the information in the common data model. If we describe a task order, all the necessary tasks must be in the "view" of this task order, however, the tasks themselves have to be modelled using ACTION and TASK within the C2IEDM and associated with the CONCEPT which models the plan to avoid ambiguity. Otherwise we would have normal tasks and task-order tasks as different concepts describing the same thing, which would be against the basic rule of data management.

methods of data modeling were used to define a common ontology is beside the point. C2IEDM is an applicable reference model for model-based data management necessary for collaboration and interoperability above the syntactic level and will play a crucial role in future systems.

C2IEDM is just one aspect within the XBML idea, namely how to represent doctrinal knowledge captured within the participating components, but it is a crucial one. How to tie this to the semantic interpretation of underlying doctrine and knowledge and what common protocols to use have been only partly dealt with in this paper. The interested reader is referred to [7] and [15] for additional ideas. The author hopes that NATO will adopt the XBML idea – and, in particular, the use of C2IEDM as a common reference model – for their needs based on the technology sketched in this paper.

Acknowledgments

The research work underlying this paper was funded by the Defense Modeling and Simulation Office (DMSO) and General Dynamics – Advanced Information Systems (GD/AIS). In addition, in particular the following individuals contributed to this paper by comments, discussion, edits, and other information exchanges means: Francisco Loaiza, Gene Simaitis and Steven Wartik (IDA), Sai Kovvuri, James Muguira and Robert King (VMASC), Bill Sudnikovich and Verlynda Dobbs (ACS), Michael Hieb (Alion), and Frank Carr (Mitre).

6. References

The following books, papers, and articles are referenced in this paper:

[1] Andreas Tolk: "Common Data Administration, Data Management, and Data Alignment Systems as a Necessary Requirement for Coupling C4ISR Systems and M&S", accepted for publication in *Information and Security, Volume 12,* 2004

[2] Andreas Tolk: "Composable Mission Spaces and M&S Repositories – Applicability of Open Standards," Spring Simulation Interoperability Workshop 2004, Paper 04S-SIW-009, Washington, D.C., April 2004

[3] Eric Newcomer: "Understanding Web Services – XML, WSDL, SOAP, and UDDI," Addison-Wesley, 2002

[4] S. Spaccapietra, C. Parent, Y. Dupont: "Model Independent Assertions for Integration of Heterogeneous Schemas," Very Large Database (VLDB) Journal, Vol. 1 (1): Pages 81-126, 1992

[5] Hong Su, Harumi Kuno, Elke A. Rundensteiner: "Automating the Transformation of XML Documents," 3rd International Workshop on Web Information and Data Management, Atlanta, GE, 2001

[6] A. Boukottaya, C. Vanoirbeek, F. Paganelli, O. Abou Khaled: "Automatic XML documents Transformations: A conceptual modeling based approach," 1st Asia-Pacific Conference on Conceptual Modelling, ACSW 2004, Dunedin, New Zealand, January 2004

[7] M. Hieb, M. Pullen, B. Sudnikovich, A. Tolk: "Developing Battle Management Language into a Web Service," Spring Simulation Interoperability Workshop 2004, Paper 04S-SIW-113, Washington, D.C., April 2004

[8] Paul Benjamin Lowry: "XML Data Mediation and Collaboration: A Proposal Comprehensive Architecture and Query Requirements for Using XML to Mediate Heterogeneous Data Sources and Targets," 34th Hawaii International Conference on System Sciences, Mauii, Hawaii, January, 2001

[9] NATO ADatP-32, Edition 2.0: "The Land C2 Information Exchange Data Model," NATO HQ, Brussels, March 2000

[10] The Multilateral Interoperability Program (MIP) website: http://www.mip-site.org <last visited March 2004>

[11] Bernd Zimmermann: "Integrated Army Modelling and Simulation Data Network," Paper #08, NATO Report RTO-MP-123: *"C3I and M&S Interoperability,"* Antalya, Turkey, October 2003

[12] U.S. Department of Defense: "DoD Metadata Registry and Clearinghouse," online specification at http://xml.dod.mil/mdreg/user/index.cfm

[13] Lionel Khimeche, Patrick de Champs: "APLET - Courses of Action Analysis and C4I-Simulation Interoperability," Fall Simulation Interoperability Workshop 2003, Paper 03F-SIW-028, Orlando, FL, September 2003

[14] Andreas Tolk, James A. Muguira: "The Levels of Conceptual Interoperability Model (LCIM)," Fall Simulation Interoperability Workshop 2003, Paper 03F-SIW-007, Orlando, FL, September 2003

[15] M. Hieb, M. Pullen, B. Sudnikovich, A. Tolk: "Developing Extensible Battle Management Language to Enable Coalition Interoperability," European Simulation Interoperability Workshop 2004, Paper 04E-SIW-064, Edinburgh, Scotland, June 2004

Authors' Biography

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 has over 12 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 related application domains, such as C4ISR or web-based services, in particular based on open standards.