MODEL-BASED ALIGNMENT AND ORCHESTRATION OF HETEROGENEOUS HOMELAND SECURITY APPLICATIONS ENABLING COMPOSITION OF SYSTEM OF SYSTEMS

Andreas Tolk

Frank Batten College of Engineering and Technology
Old Dominion University
Norfolk, VA 23529, U.S.A.

Charles Turnitsa
Saikou Diallo

Virginia Modeling, Analysis and Simulation Center
Old Dominion University
Norfolk, VA 23529, U.S.A.

ABSTRACT

One of the main challenges for Homeland Security applications is the fact that the different supporting organizations, services, and nations come to the table with existing information technology, supporting established business and organization processes, and using organization-specific data models. This paper shows how to support multi-organization processes with a federation of their heterogeneous IT-solutions based on the alignment and orchestration of applications with regard to the underlying models of those solutions. While processes are orchestrated and aligned top-down, the supporting IT is migrated into a Homeland Security System-of-Systems bottom-up. Web services allow the loose coupling of participating systems and the consistent application of data engineering allows the auto-configuration of data mediation layers. This is made possible by considering first the solutions themselves, and their models (the top-down approach), and only then the application of data engineering to aligning those models (the bottom-up approach).

1 INTRODUCTION

The U.S. Department of Homeland Security Strategic Plan (U.S. Department of Homeland Security 2004) identifies six guiding principles in support of the philosophy that informs and shapes decision-making and provides normative criteria that govern the actions of policy makers and employees in performing their work in this domain. The guidelines are: protecting civil rights and civil liberties, integrating the actions of the 22 previously disparate agencies, building coalitions and partnerships, developing human capital, and innovation. Of these six guiding principles, "integrating the actions" and "building coalitions and partnerships" present challenging requirements for migration and interoperation because each of the 22 previously disparate agencies were in existence before Homeland Security established new requirements. Each

agency comes with its own policies and processes, proprietary IT systems supporting these processes, and data models structuring the necessary information within applications. In order to support the mission of Homeland Security, these stove-piped solutions must be migrated to a common technical standard and integrated to support the new processes, many of which were not foreseen when the existing processes were defined and the supporting IT systems procured. Buying new systems for all participating agencies is out of question, as it is too expensive. Furthermore, employees are already well trained and familiar with their systems.

In general, the principle of building coalitions and partnerships adds additional layers of complexity. First, it is nearly impossible to foresee whom coalition partner in future operations will be. Hence, it is hard to define all underlying assumptions and constraints for Homeland Security processes and supporting systems. The systems must be configurable and open in order to support new tasks and requirements. Furthermore, no single agency can be responsible in a consortium of international partners. Nonetheless, the first goal of Homeland Security applications documented in (U.S. Department of Homeland Security 2004) is awareness, which means to identify and understand threats, assess vulnerabilities, determine potential impacts and disseminate timely information to our homeland security partners and the (American) public. In other words, a roadmap is needed to capture the differences and agreements, and support migration towards the loose coupling of necessary systems. This roadmap will enable collaboration on the short term, the identification of information (process and data) representation needs within the system-of-systems on the mid-term, and the alignment of the collaboration processes to increase efficiency on the long term. For the supporting IT systems, the goal must be homogeneous support of common business process based on heterogeneous IT systems.

The authors have several years of experience on such topics from applications in the military domain. The chal-

lenges are quite similar when the operational ideas and doctrines of different participating nations, e.g. in NATO operations, or different military services, e.g. aligning Army, Air Force, and Navy requirements for the common use of air space in joint operations, must be aligned and orchestrated. Services and nations are also equipped with IT systems that generally were not designed to support such collaboration. Current military operations involve various allies and partners coming from quite different domains like technical support, disaster relief, commercial tele-communication, and many more. The ability to create one common picture to coordinate the operations becomes crucial. As Larry Wentz already pointed out in his work on Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR) systems during the operations in Bosnia (Wentz 1998): *"Although integrated C4ISR services are the desired objective, the realities tend to drive the solution to stove-piped implementations. In spite of technology advances, this will likely be the case for some time to come. There will continue to be uneven C4ISR capabilities among coalition members who will continue to rely on systems with which they are most comfortable – their own."* The situation within Homeland Security is, if anything, more complex due to the misalignment of training, resources, and goals of the 22 individual organizations brought together under the U.S. Department of Homeland Security's umbrella.

This paper summarizes the main ideas on migration and integration principles applicable to Homeland Security applications based on open and web-based standards using a model-based approach. While doctrine and business processes are aligned top-down resulting in common processes and information exchange requirements, the underlying IT systems are aligned bottom-up based on a rigorous application of data engineering principles. This article will summarize the ideas of model-based alignment and orchestration of business process, which build the conceptual prerequisites for effective data engineering, before describing the data engineering process itself. The general focus is data engineering, but the paper as a whole gives an overview for a roadmap on how to integrate Homeland Security applications based on open and web-based standards. The feasibility and applicability of these ideas has been proven for the military domain. The authors see no reason why this should not be possible for Homeland Security.

2 MODEL BASED ALIGNMENT AND ORCHESTRATION

This section summarizes the conceptual prerequisites for efficient data engineering. Before the data engineering processes, which are in the focus of this paper, can be applied bottom up, some analysis based on the ideas of system dynamics must be performed to derive top down the processes that have to be aligned and the information that

has to be exchanged. In order to support the alignment of processes, supporting IT systems, implementations, data exchange requirements, and data models, a top-down model driven approach for the business models and a bottom-up model driven approach for data is recommended. The NATO Code of Best Practice for Command and Control Assessment (NATO 2002) drives these ideas. It was developed to align the different national requirements and constraints in the light of aligned Command and Control for international operations under the umbrella of NATO. These operations are no longer limited to the traditional use of military forces. Alignment is needed with distributed teams including military, coalition and other Non-Government Organization (NGO) actors operating under stress and their varying decision-making behaviors. Particular interest must be paid to behavior of and interaction with non-military organizations, political groups, and amorphous groups such as crowds and refugees. The requirements for Command and Control in these environments are clearly relevant for Homeland Security applications.

The Modeling and Simulation (M&S) community distinguishes between models and simulations. *Models* are a purposeful abstraction of reality, capturing constraints and assumptions resulting in a conceptualization of the problem to be solved, the environment in which it has to be solved, and relevant actors, their behavior, and the relationships of interest to solve the problem. *Simulations* are implementations of models executable over time, which means that they allow what-if-analysis, evaluation of alternative approaches, etc. The NATO Code of Best Practice emphasizes the importance of models to gain a common understanding of the problem to be solved and the alternative solution strategies. Based on the general assumptions, the following advantages of model driven approaches to reach alignment of heterogeneous decision-making processes have been published in Tolk (2003). They should be the basis for Homeland Security since inter-agency and international interoperability are issues beyond technical interoperability:

- *Models help the decision makers understand the key mechanisms of an existing process.* A model provides a clear picture of acting entities, roles, relations, and tasks. This is needed to understand the processes of the allies as well as the processes of the non-military partners and vice versa.
- *Models act as the basis for creating suitable information systems that support the process.* The model comprises descriptions of process that can be used to identify necessary support. Furthermore, the sub-processes already supported by IT in the various participating organizations are displayed. This includes systems' interfaces as well as their information capability that is available in-

formation that can be delivered to other systems as well as suitable information that can be computed to deliver new insights. Therefore, the model puts the various existing systems into their place within the federated system of systems supporting the overarching operations and also serves as the requirement driver for additional IT support.

- *Models can be used to improve the current structure and operation.* By creating a common description of the overall operation, participating organizations and supporting systems, redundancies as well as bottlenecks become obvious. Necessary changes can be identified and solutions can be derived and agreed on based on a common model.

- *Models show the structure of innovated solutions.* The model becomes the basis for a common action plan supporting radical as well as incremental changes. The desired end state and the necessary steps leading from the status quo to this end state are part of the model. The model itself becomes an important management instrument that orchestrates the necessary improvements in parallel and distributed events.

- *Models can serve as a basis to evaluate new ideas.* Models can be used to copy other structures, and evaluate processes used by other partners – or opponents – in the environment in which the operation takes place. As the model comprises the necessary detail needed to derive a conceptual or functional model of the mission space, support by M&S directly becomes possible. Respective experiments can help evaluate such future concepts. An appropriate model can be used to orchestrate respective efforts and helps create a common understanding of all participating institutions.

- *Models facilitate the identification of potential re-use of existing solutions.* Although every operation is special and unique, many processes are supported by standard solutions. Additionally, when using a common model, the identification of processes supported in other operations and that can be modified easily to support the current effort becomes feasible with minimal effort.

The top-down use of models in understanding the alignment of different processes must be accompanied by an analysis of required information exchange between the processes resulting in data exchange requirements between the underlying IT systems. The advent of the Extensible Markup Language (XML) and web services marks a significant starting point for a new era of interoperation. However, while XML enables separation of data definition a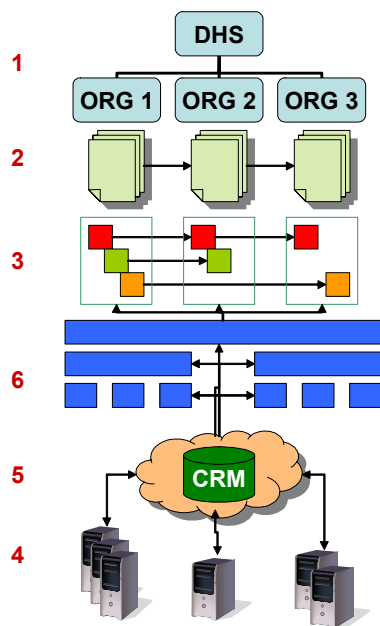nd data content, it doesn't ensure that data exchanged is interpreted correctly by interoperating systems. This motivates the consistent application of Data Management to support the unambiguous definition of data elements for information exchange. Using a common reference model improves this process leading to "Model Based Data Engineering (MBDE)" (Tolk and Diallo 2005). The results can be used immediately to configure data mediation layers integrating services into an overall Service Oriented Architecture (SOA).

As already mentioned, the alignment of the operational goals of the homeland security organizations with the IT systems supporting those organizations is accomplished by applying both a top-down and also a bottom-up approach. The top-down approach concerns itself with deriving accurate conceptual models of the operational goals, and then revealing the important data concepts that constitute those models – this approach will be described here, and is important as it generates a goal for the bottom-up approach. The bottom-up approach applies MBDE principles resulting in a common reference model (CRM) describing the information exchange that needs to be supported between the systems. These must be the same concepts that are the results of the top-down approach. The bottom-up approach, with its emphasis on data engineering, is described in detail in the next section, and is the main focus of the research contributions supporting this paper.

To begin the top-down approach, the first step entails identification of the organizations that will participate in satisfying a particular operation. This involves not only each organization, but also an understanding of what each organization is contributing to the operation, as well as what systems the organization has to support that contribution. This is illustrated in Figure 1 as the first step in the top-down approach.

The second step in the top-down approach is to construct a conceptual model of how each of the contributing organizations will make their contribution to the operation being discussed. Such a model will be based upon the doctrine of the contributing organizations. This model is based on the different modeling strengths described above, as it can result in not only a picture of what is expected to happen, but also provides a basis for showing how the different processes will interact with each other. This model is a conceptual model, described in Robinson (2006) as *"a non-software specific description of the simulation model that is to be developed, describing the objectives, inputs, outputs, content, assumptions and simplifications of the model."* This definition serves very well what should be accomplished at this step, identifying and describing objectives, inputs, outputs, content, assumptions, and simplifications.

The third step is the identification of information exchange events between the processes. The second step resulted in conceptual models of all processes supported by each participating organization. The new common

Figure 1: Top-down, bottom-up approach.

processes in support of Homeland Security build the basis for the third step. The analyst identifies what processes of the participating organizations can be used to be composed in support of the Homeland Security applications. In this process, he also identifies the conceptual (data) domains and data element concepts needed to describe the information exchange necessary between the processes on the conceptual level. The result of the top down approach is the conceptual understanding what information exchanges occur when, between which processes, what are the business objects into which the atomic information elements are composed or aggregated, and which organizations – and hence which supporting IT systems – contribute as source or target systems to this system-of-systems supporting Homeland Security applications.

The bottom up approach, with its emphasis on Data Engineering is discussed in brief in the following section. The top-down approach is important, as it results in an aligned conceptual model of processes and information exchange events, and these results are building the conceptual prerequisites for processes of the bottom-up approach.

3   TOP DOWN DATA MODELING

Data engineering, such as defined in Tolk (2004), uses data modeling artifacts describing the information exchange capabilities of interfaces between systems supporting interoperability and composability of applications that collaborate by exchanging information between each other to support a common process, often describable as a work-

flow. As previously mentioned, the U.S. Department of Homeland Security is a complex organization, formed from bringing together a number of previously discrete organizations. Each of these organizations brings with it a number of different technological solutions for the representation, consumption, and production of data. Data engineering supporting such a complex task has many facets that are supported by different standards. These standards need to be applied and executed in an orchestrated way by the a leading organization first, and then propagated to all subordinate components of that organization. In order to accomplish the data engineering tasks described in section 4 below, first a data model must exist describing each of the data needs, the data themselves, and the business rules of the organizations employing the data. The first category of challenges is to identify the processes/workflows and supporting information infrastructure and evaluates questions such as:

- Which processes are conducted in the headquarters and subordinated commands must be identified and how can they become operationally aligned?
- Which systems are applied in support of these processes and which system components are providing the technical functionality required to support them?
- Which information exchange requirements can be derived from the operational process?

- Which information exchange capabilities of the supporting systems and system components are available?

This analysis driven by operational needs leads to the identification of operationally relevant information infrastructure components; more technical questions can be addressed, such as:

- Which data sources are important for the task?
- What are the characteristics of these data sources?
- Are there any constraints the users must be aware of, such as security concerns, reliability of the data sources, costs to obtain the data, etc.?

These questions regarding the data sources must be completed with information regarding the data, such as its structure, and its meaning:

- What is the resolution of the data?
- What is the scope of the data?
- How is the data structured/modeled?
- What is the meaning of the modeled data entities?

As the data sources will be heterogeneous, the data must be mapped to a common reference model accepted by the supported community of interest. To support the necessary data mediation from data sources, the data and data source descriptions need to be described in a common way. This is done using "data about data," which are metadata.

Data modeling addresses all these needs for information exchange between systems in support of business/operational processes in and between the headquarters and the commands and test centers, these information exchanges respective data sources, metadata describing content and structure of these data sources, and metadata describing content and structure of comprised data. As such, it is captured by respective artifacts, which are preferably produced following applicable standards; what operational activities in and between the headquarter exists, which systems provided the necessary functionality, what data needs to be exchange between these systems, what data exist in these systems (plus pedigree, reliability, constraints, etc.), how these data are structured, and what exactly these data mean. The recommended common data engineering architecture is build on the foundation of established handbooks, guidance, best practices, manuals, and other established procedures.

In order to validate this method as a practice to be followed, one need look no further than the many successes that adherents of the Model Driven Architecture (MDA) have claimed. What results from the top-down approach described here is expressively equivalent to the platform independent model (PIM). The PIM is a collection of modeling efforts that identify actors, processes, and model ob-jects that are the tokens of those processes. The same is derived from the above described top-down approach, the difference being in the step-wise approach described here. MDA suggests that some Model Transformation Language (MTL) exists or can be formulated (and several have been presented and marketed commercially) that will then translate the PIM into a Platform Specific Model (PSM), which serves as the framework for operational software design. As we are dealing with already existing, heterogeneous data models that step will not apply to the bottom up approach described below, yet the results (again) will be similar. For additional details on the MDA, see numerous documents and reference available at the website of its parent organization (OMG 2007).

4 BOTTOM UP DATA ENGINEERING

Data Engineering is a framework that guides information exchange between all participating organizations, and therefore their supporting IT systems. Commercial solutions usually generate XML interfaces on top of existing data interchange solutions. However, XML interfaces thus generated only ensure the exchange of bits and bytes at the technical level. In order to exchange information and ultimately knowledge, and ensure the interpretation of data in a consistent way leading to the same information, knowledge, and ultimately awareness within the services and their users, each service 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 services composed into a distributed application within the overall system. To generate the answers to these questions is the objective of data administration, data management, data alignment, and data transformation, which are the building blocks of a new role in the interoperability process: *Data Engineering* (Tolk 2003a).

4.1 Traditional Data Engineering

The composing terms of Data Engineering are defined in (Tolk 2003a) as follows:

- *Data Administration* is the process of managing the information exchange needs that exist between the services, including the documentation of the source, format, context of validity, and fidelity and credibility of the data. Data Administration therefore is part of the overall information management process for the service architecture. Data Administrations must answer the questions: "Where are the data? In what format? How can the data be accessed?"
- *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. Data are described by propertied concepts describing the universe of discourse as well as their relations. Data Management answers the question: "What do the data mean?"

- *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 available data, e.g., using the means of aggregation or disaggregation. Data Alignment answers the question: "Can all needed data be obtained?"
- *Data Transformation* is the technical process of aggregation and/or disaggregation of the information entities of the embedding systems to match the information exchange requirements including the adjustment of data formats as needed. Data Transformation answers the question: "How to transform/mediate data?"

Within XML environments, data management really becomes tag set management. The challenges are not trivial. They are closely related to problems to be solved in heterogeneous, distributed database environments. In their works on heterogeneous data federations, Spaccapietra and his colleagues identify the following four classes of conflicts to be solved by data management (Spaccapietra, Parent, and Dupont 2002, Parent and Spaccapietra 1998). It is worth noting that are applicable to semantic XML tag set management as well:

- *Semantic Conflicts* occur when concepts of the different local 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* describe homonyms, synonyms, and different names for the same concept, different attributes or slot values for the same concept, etc.
- *Heterogeneous Conflicts* result from substantially different methodologies being used to describe the concepts.
- *Structural Conflicts* results from the use of different structures describing the same concept.

Spaccapietra et al. concluded that a generic meta data model comprising only objects and attributes for values and references is needed to support efficient data management. Their model can easily be mapped to XSLT supported structures.

The resulting description comprises data describing the conceptual domain (what concepts are modeled, what data element concepts are needed to describe these concepts) as well implementation specifics (what data elements are used in the implementation, what is the value domain chosen for their implementation). This allows to apply the metadata registry standard ISO/IEC 11179 (Part III) as described by Winters, Gorman, and Tolk (2006). The same ideas are also applied in the recommended common data engineering architecture for the U.S. Army Test and Evaluation Command (Tolk and Aaron 2007). An application is described by Perme et al. (2007).

4.2 Model-based Data Engineering

Model-based Data Engineering introduces the idea of using a Common Reference Model (CRM) for data management (Tolk and Diallo 2005). A CRM captures the meaning of data and their relations. In other words, the CRM manages the Conceptual Domain and the Data Element Concepts as defined in ISO/IEC 11179 for the common information sphere of the aligned decision-making processes of participating organizations. The other steps of data engineering are not changed, but they are supported by the CRM as well. When a CRM is used, data alignment becomes a simple comparison of the mapping results. Data management maps information elements of the source and the target to the CRM. After target and source are mapped to the CRM, we can compare the mapping of the source model to the CRM with the mapping of the target model to the CRM. If every information element of the source and target model is mapped to an element of the CRM, the models are aligned. As data administration and data transformation are already well supported by web-based standards, data management becomes the dominant challenge.

Model Based Data Engineering processes go beyond simply mapping attributes and tables to each other or creating an interface with some translation technology applied to it. Starting with a core model of the CRM, the continuous application of data management perpetually enhances format and content represented (Tolk and Diallo 2005). For this continuous improvement, two ideas are significant:

- Every time a model of higher resolution is made interoperable (i.e., via a mapping) with the core model, the core model's resolution necessarily increases. Typical examples are adding enumerations within the applicable property values or more details – mostly modeled in the form of additional tables within a view – to describe a higher concept.
- Every time a new concept not in the core model is mapped, this new concept is integrated. In particular when merging models from complementary domains – such as army, air force, and navy models in the military area – this happens quite often. In Homeland Security applications, this will be more often the rule than the exception in the initial phase, but the process ensures that common

concepts and individual concepts– which are as important – unique to a participating organization are identified.

Following this process, data engineering gradually refines, enhances, and extends the CRM from the initially agreed-upon core model. As a result the CRM continuously grows in its applicability to the applications summarized in the supported domain. It is critical that the extensions and enhancements of the CRM follow strict rules to ensure that earlier mappings remain valid. A data management group responsible for these activities is needed in bigger organizations. An example is the NATO data administration group (NDAG) of the Multilateral Interoperability Programme (MIP) of NATO. Zimmermann (2003) describes how these structures can be used to set up a national M&S solution.

In order to ensure not only the theoretic availability of information with data alignment and its accessibility by demanding systems in a timely fashion, business objects must be defined and managed. Arpinar and colleagues use such compositions to define the elements supported by their ontology, which reflects the application domain (Arpinar et al. 2005). Srivastava and Koehler identify the need for orchestration using the information flow underlying the supported business process as well (Srivastava and Koehler 2003). The authors apply these ideas as well. Using the documentation generated from the alignment of business processes between the participating organizations the temporal and causal relations of information exchange elements are captured. The use of data engineering unambiguously identifies the participating concepts and underlying data element concepts. In summary, the combination of business processes and data engineering results in the unambiguous of concepts and the context in which they are exchanged.

Many application domains already have a sort of CRM. If this is not the case, the federated database approach as described by (Sheth and Larson 1990) can be applied to generate the core of a CRM based on the information exchange events captured in the top down processes.

4.3 Result of Data Engineering

The results of data engineering are directly usable to configure XML based data mediation services (Tolk 2004). The idea is that first the top-down analyses identifies business services that are used. Based on the alignment of such services, information exchange events between the IT systems of the participating organizations are defined. These information exchange events are defined on the conceptual level and result in the definition of concepts whose descriptions need to be exchanged between the systems. Data engineering is then conducted bottom up for the participat-

ing systems and can be limited to the necessary concepts. While it is theoretically possible to align all concepts of target and source systems, practically it seldom makes sense to align concepts that are unique to one system.

The CRM comprises all concepts, as captured in the conceptual domain, and all defining properties, as captured in the data element concepts, as well as their relations and rules. Furthermore, each concept models a piece of information that on its own is already of value for participating applications. We also request that the information in the CRM be "atomic" for the participating systems, which means no participating application splits the information into two or more implementing data elements. All necessary views of the supporting IT systems are generated from a composition or aggregation of atomic pieces. Composition means that data for an application can be generated from the CRM by composing the necessary data element concepts. Aggregation means that the data obtainable from the CRM must be modified before they fulfill the requirements of the application. In other words: If the CRM concept is represented in the participating system, composition is sufficient; if additional constraints of the underlying CRM must be taken into account, or if some information must be modified using functions to satisfy the requirements of the application, aggregation is necessary. This additional knowledge can be captured in relations, which are needed to enable compositions, and rules and processes, which will result in functions. The *data mediation services* use the correlating three types of information services based on these ideas:

For each concept, there is one service allowing inserting, updating, and selecting information. These are *atomic services* directly accessing the concepts captured in the CRM. The services are transparent to the user so that individual concept access is supported.

For higher objects, which are defined as a collection of information of interest distributed over more than one concept, views of the CRM are defined based on concepts and relations. These views are presented as one service, but they make use of the underlying atomic services. They are called *composite services*. The application must ensure that underlying rules and processes required by the CRM are followed. Composite services support the rapid integration of new models; however, the user is responsible for the integrity of the underlying data as composite services only retrieve and store information based on atomic services.

For business objects, in particular for those identified in the information exchange requests, access routines are defined to not only access the necessary information, but also ensure that underlying rules and processes required by the CRM are followed and aggregation of information is conducted. These are *aggregate services*, which support data integrity as well as accessibility of data. They use concepts, relations, and rules.

Finally, these services have been prototypically implemented as web services. Web services are discrete web-based applications that interact dynamically with other web applications. The fundamental idea behind web services is the integration of software applications as services using a defined set of industry-supported, open standard technologies that work together to facilitate interoperability between heterogeneous systems, either within an organization or across the Internet. Because they are based on standard interfaces, they can communicate even if they are running on different operating systems and are written in different languages. They are widely supported by industry and already successfully applied in a wide range of different domains. If the atomic, composite, and aggregate services are implemented as web services, they can be registered within a registry supporting Homeland Security applications than can be invoked dynamically, which eliminates the need to preplan the invocation sequence. Applications can be integrated into the system-of-systems "on the fly," as long as the management process as described before is captured by the necessary metadata and all services are registered.

4.4 Validation of the Bottom-Up Approach

The application of the bottom-up approach described here has been shown to be successful in several projects that the authors, and others, have been involved on. One of the most successful, and indeed still ongoing, applications has been for the Joint Event Data Initialization Services (JEDIS) project for the Joint Advanced Training and Technology Lab (JATTL). In this project, the exchange of data for a common purpose by heterogeneous organizations, employing heterogeneous data models for origination of data, was accomplished by applying first the top-down and then the bottom-up approaches described in this paper. The organizations (and their systems) who were going to be the final users of the data formed the highest level of the top-down approach, then their business practices were revealed, and finally the data needs for those practices were specified, and found to exist (as composable elements) in a common data model.

Once this top-down approach was completed, then the bottom up approach described in sections 4.1-4.3, above, was employed. The originator of the architecture was the Virginia Modeling Analysis and Simulation Center, yet the final developer who performed the bulk of the operational implementation was Gestalt LLC. Details of the operation can be found in Perme et al. (2007).

5 ENABLING COMPOSABILITY

Both approaches are necessary to enable composability. Supporting only one approach will lead to variances, conceptual gaps, and insufficient implementations.

- The top-down analyses leads to the identification of common processes, supporting IT systems in the heterogeneous IT infrastructure, and the supporting applications, including M&S applications. It results in a common conceptual understanding of the overall process, what role the systems to be integrated will play in this process, and what conceptual alignments are needed to avoid logical and conceptual mismatches in the assumptions and constraints.
- The bottom-up approach ensures that application and implementation constraints of the participating systems and services are captured and documented based on engineering methods. The result is the definition of standardized information exchange specifications and configurable mediation layers as described by Tolk (2004).

Both approaches together ensure that collaboration of the systems is based on a common model and the integration of systems into a system of systems is based on engineering principles ensuring the exchange of information in the right format at the right time based on the concept of operations derived from aligning the participating business models and support IT solutions.

It should be pointed out that the alignment of conceptual ideas is increasingly important in the Homeland Security domain. While in the military domain, warfighters share a common understanding based on centuries of experiences on how to conduct battles against or with each other, we are now facing organizations that do not necessarily share such a common background. To capture constraints and assumption and the applied abstraction of reality becomes essential to avoid misalignments and resulting misunderstandings and bad decisions.

ACKNOWLEDGMENTS

REFERENCES

Arpinar, B., R. Zhang, B. Aleman-Meza, and A. Maduko. 2005. Ontology-driven Web services composition platform. *Journal on Information Systems and E-Business Management* 3 (2): 175-199.

International Standard ISO/IEC 11179. 2004. Information technology – metadata registries (MDR), part III. ISO Press.

North Atlantic Treaty Organization (NATO). 2002. *NATO code of best practice for command and control assessment.* Revised ed. Washington, DC: CCRP Press.

Object Management Group (OMG). 2007. Available via <http://www.omg.org/mda/> [accessed July 1, 2007].

Parent, C. and S. Spaccapietra. 1998. Issues and approaches of database integration. *Communications of the ACM* 41(5): 166-178.

Perme, D., S. Diallo, C. Pandolfo, B. Tedesco, H. Tran, and A. Tolk. 2007. Joint event data initialization services (JEDIS) – implementing a service oriented architecture for initialization. *IEEE Spring Simulation Interoperability Workshop*. IEEE CS Press.

Robinson, S. 2006. Issues in conceptual modelling for simulation: setting a research agenda. In *Proceedings of the 2006 OR Society Simulation Workshop*.

Sheth, A. P. and J. A. Larson. 1990. Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys* 22(3): 183-236.

Spaccapietra, S., S. Parent, and Y. Dupont. 1992. Model independent assertions for integration of heterogeneous schemas. *Very Large Database (VLDB) Journal* 1(1): 81-126.

Srivastava, B. and J. Koehler, J. 2003. Web service composition – current solutions and open problems. In *Proceedings ICAPS 2003 Workshop on Planning for Web Services*.

Tolk, A. 2003. Common data administration, data management, and data alignment as a necessary requirement for coupling C4ISR systems and M&S. *Information and Security* 12(2): 164–174.

Tolk, A. 2003a. Beyond technical interoperability - introducing a reference model for measures of merit for coalition interoperability. In *Proceedings of the Command and Control Research and Technology Symposium (CCRTS)*, CCRP Press.

Tolk, A. 2004. XML mediation services utilizing model based data management. In *Proceedings of the 2004 Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 1476-1484. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Tolk, A. and R. D. Aaron. 2007. A common data engineering architecture. White Paper. Army Test and Evaluation Command, Washington, DC.

Tolk, A. and S. Y. Diallo. 2005. Model based data engineering for Web services. *IEEE Internet Computing* 9(4): 65-70.

Tolk, A., S. Y. Diallo, C. D. Turnitsa, and L. S. Winters. 2006. Composable M&S Web services for netcentric applications. *Journal for Defense Modeling and Simulation* 3(1): 27-44.

U.S. Department of Homeland Security 2004. Securing our homeland, U. S. Department of Homeland Security strategic plan. Washington, DC.

Wentz, L. 1998. Lessons from Bosnia: The IFOR experience. National Defense University, Institute for National Strategic Studies, NDU Press: Vienna, VA.

Winters, L. S., M. M. Gorman, and A. Tolk. 2006. Next generation data interoperability: it's all about metadata. *IEEE Fall Simulation Interoperability Workshop*. IEEE CS Press.

Zimmermann, B. 2003. Integrated Army modelling and simulation data network. In *Proceedings of the NATO M&S Group Conference on C3I and Modelling and Simulation Interoperability*. NATO Document No. RTO-MP-123.

AUTHOR BIOGRAPHIES

ANDREAS TOLK is Associate Professor for Engineering Management and Systems Engineering of Old Dominion Universities Modeling, Simulation, and Visualization Faculty. He is also a Senior Research Scientist at the Virginia Modeling Analysis and Simulation Center (VMASC). He holds a M.S. in Computer Science (1988) and a Ph.D. in Computer Science and Applied Operations Research (1995), both from the University of the Federal Armed Forces of Germany in Munich. He is a member of SCS and SISO. His e-mail address is <atolk@odu.edu>.

CHARLES D. TURNITSA is a Ph.D. candidate at the Virginia Modeling Analysis and Simulation Center (VMASC) of the Old Dominion University (ODU). He received his B.S. in Computer Science (1991) from Christopher Newport University (Newport News, Virginia), and his M.S. in Modeling & Simulation (2006) from ODU. His Ph.D. research under Tolk focuses on the domain of dynamic and fractal ontology models for M&S interoperability. His e-mail address is <cturnits@odu.edu>.

SAIKOU Y. DIALLO is a Ph.D. candidate at the Virginia Modeling Analysis and Simulation Center (VMASC) of the Old Dominion University (ODU). He received his B.S. in Computer Engineering (2003) and his M.S. in Modeling & Simulation (2006) from ODU. His Ph.D. research under Tolk focuses on the domain of Model Based Data Engineering and web services for M&S applications. His e-mail address is <sdiallo@odu.edu>.