MDA Journal

July 2004

David S. Frankel
David Frankel Consulting

df@DavidFrankelConsulting.com

The computer industry is large, with many streams of thought and practice. Differing experiences, assumptions, and fundamental concepts sometimes stymie the potential for synergy among different communities. Sometimes we labor without being aware of the significance of other people's work to our own endeavors.

I was only peripherally aware of the Semantic Web until last year, when Erick and Linda Von Schweber of Synsyta LLC brought together some MDA experts and authorities in Knowledge Representation and the Semantic Web for a mini conference in northern California. By the time the conference ended, I had had an "aha experience," during which I began to understand the potential for MDA and the Semantic Web to work together.

Previous to this event, the only thought I had as to the potential for integrating the Semantic Web and MDA was the idea—a rather obvious one to MDA aficionados—that MOF metamodels of the Semantic Web languages would help to integrate ontologies into the MDA world. I did not appreciate the role that reasoning could play in making MDA more scalable. Erick and Linda's initiative had awakened me. They have been doing some industrial work that is exploring such possibilities and also have been looking at opportunities for MDA to help ease the introduction of the Semantic Web into the enterprise and embedded software arena.

A few months later, Elisa Kendall, CEO of Sandpiper Software in Los Altos, California, approached me. Her company makes ontology tooling, and is one of the submitters to the OMG's Ontology Definition Metamodel (ODM) RFP. She was way ahead of me. She understood the promise of combining these technologies and the importance of it to her company, and was willing to put money on the table to involve me in the ODM activities, which she saw as crucial groundwork.

Thus, I joined the already very strong ODM submission team, which also includes Louis Hart and Patrick Emery of ATT, Bob Colomb of DSTC, Marko Boger of Gentleware, and Dan Chang of IBM. Elisa arranged for Deborah McGuinness and Pat Hayes, who are key authors of the Semantic Web specifications, to help as well.

In this month's MDA Journal, Elisa and I join with Deborah and Pat to explain the potential for synergy between MDA and the Semantic Web, and to discuss the direction that the ODM work is taking. Explaining the basic business case is just the first step. Much work remains…

Until September,

David Frankel

MDA Journal

July 2004

David S. Frankel

Patrick Hayes

Elisa F. Kendall

Deborah L McGuinness

A Model-Driven Semantic Web

Reinforcing Complementary Strengths

Table of Contents

EXECUTIVE SUMMARY

The OMG's Model Driven Architecture® (MDA) initiative and the W3C's Semantic Web project initially proceeded in parallel with no coordination. Recently, however, some of the key people from both communities have been collaborating to bridge these two technologies. We have come to understand that, while there is some overlap between the two, each has unique strengths that can benefit the other.

Therefore, we have joined in a project, sponsored by the Object Management Group (OMG), to define standards for bridging between these technologies.

This article explains the potential for synergy and provides an overview of the technical work in progress toward the goal of a "Model-Driven Semantic Web."

QUICK OVERVIEW OF MDA®

The two key MDA standards are UML® and MOF™. UML is better known than MOF, but MOF is the most fundamental MDA technology.

Quick Overview of MOF™

The OMG's Meta Object Facility (MOF™) defines the metadata architecture that lies at the heart of MDA. MOF technology helps to automate metadata management. What the MDA world calls metadata includes database schema, UML models, workflow models, business process models, business rules, API definitions, configuration and deployment descriptors, and so on.

MDA seeks to deal with the current reality in which each kind of metadata is an island that is disconnected from other metadata. Toward this end, MOF defines standards for *automating the physical management and integration of different kinds of metadata*. To add a kind of metadata to the kinds of metadata that MDA tools can manage, it is necessary to define a MOF model of that kind of metadata. Such a model is usually called a *metamodel*.

MOF-based tools use metamodels to generate code that manages metadata embodied as XML documents, Java™ objects, and CORBA® objects. The generated code presents and implements Java or CORBA APIs that make it possible to read and manipulate the metadata, and also serializes metadata from Java or CORBA objects to XML documents and vice versa. By automating the production of such code according to standardized patterns, MOF tools relieve programmers of tedious, repetitive chores and enforce a consistent approach to managing disparate kinds of metadata.

MOF is actually a core standard that is augmented by the following adjunct standards[1]:

- *XML Metadata Interchange (XMI®)*, which defines how to manage MOF metadata as XML documents[2]
- *CORBA Metadata Interface (CMI),* which defines how to manage MOF metadata as CORBA objects[3]
- *Java Metadata Interface (JMI),* which defines how to manage MOF metadata as Java objects[4]

[1]For detailed background on MOF, XMI, CMI, and JMI, see [FRANKEL 2003], chapter 5.
[2] [XMI]
[3] [CMI]
[4] [JMI]

XMI and CMI are OMG standards, while JMI is a Sun Java Community Process standard. Some MOF-based tools also support Web service interfaces to MOF metadata.

Kinds of metadata for which the OMG has defined or is defining metamodels include relational database modeling, hierarchical database modeling, online analytical processing (OLAP), business process definition, business rules specification, XML, UML®, and CORBA IDL. As the OMG continues to define new metamodels, an infrastructure for connecting organizations' metadata islands is forming.

Notably, the Eclipse integrated development environment uses MOF technology to create an ecosystem in which multiple development tools can coexist and share metadata. Eclipse was originally an IBM open source product that now is under the control of an independent organization and has gained enormous traction in the industry over the past few years. It has extensive support for plugging third party tools into its environment, and uses technology based on MOF and XMI to integrate the metadata that drives multiple cooperating tools in its ecosystem.

The Role of UML® in MDA

UML is itself defined via MOF. In one sense, therefore, UML is one of potentially many modeling languages from the standpoint of MDA. However, there are some factors that make UML more than just another language for MDA.

1. *UML has significant penetration in the market place*. This is not to say that UML is without its detractors and faults, but it is a simple fact that many people in the software industry use it.
2. *UML is extensible, through UML profiling*. UML profiles are specialized dialects that target different kinds of systems or different aspects of systems. The OMG has standardized some UML profiles, such as a profile for modeling real time systems, one for modeling software test facilities, and one for modeling CORBA-based systems. Many tool vendors and IT shops have developed their own, proprietary UML profiles.

When an architect must define a modeling language, defining a UML profile has the advantage of making it possible for a modeler to use the language via a generic UML tool. Sometimes, however, UML profiling is too constraining, and in such cases it often makes sense to use MOF for language definition.

Thus, via UML profiling and MOF, MDA is architected from the ground up to support multiple modeling languages and methodologies.

QUICK OVERVIEW OF THE SEMANTIC WEB

The Semantic Web, as envisioned by Tim Berners-Lee and many others since, is a logical extension of the current Web that enables explicit representations of term meanings. It makes it possible to describe the Web in a precise form, thus enabling greater understandability by humans and greater processibility, by machines[5]. The W3C's Semantic Web Activity has evolved from more than

[5] [DACONTA ET AL 2003]
[6] See, for example, [BAADER ET AL, 2003]

two decades of research on languages for knowledge representation and automated reasoning[6].

The Semantic Web uses explicit descriptions of terms and their interrelationships to enable programs to "understand" the meanings of terms. In Tim Berners-Lee's words, "The first step is putting data on the web in a form that machines can naturally understand, or converting it to that form. This creates what I call a Semantic Web – a web of data that can be processed directly or indirectly by machines[7]… The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."[8]

RDF and OWL

The W3C Semantic Web Activity's RDF Core Working group has produced the Resource Description Framework (RDF)[9] and the Web Ontology working group[10] has produced the Web Ontology Language (OWL)[11] . OWL is based on RDF.

The following quotation from the abstract of the W3C's OWL Web Ontology Language Guide[12] restates the problem that OWL and the Semantic Web aim to address:

> The World Wide Web as it is currently constituted resembles a poorly mapped geography. Our insight into the documents and capabilities available are based on keyword searches, abetted by clever use of document connectivity and usage patterns. The sheer mass of this data is unmanageable without powerful tool support. In order to map this terrain more precisely, computational agents require machine-readable descriptions of the content and capabilities of Web accessible resources. These descriptions must be in addition to the human-readable versions of that information.
>
> The OWL Web Ontology Language is intended to provide a language that can be used to describe the classes and relations between them that are inherent in Web documents and applications.

OWL ontologies contain axioms and facts about Web resources, which constitute definitions and metadata about the resources. The RDF and OWL specifications use mathematical techniques to formally define the semantics of the languages, which makes it possible for specialized tools to *automate reasoning about such axioms and facts*.

OWL axioms and facts are formed from class declarations and from a set of designated constructors such as "intersectionOf," "subClassOf," "sameAs," "inverseOf," "hasValue," "allValuesFrom," and so forth. Additionally, annotation and versioning information may be captured using notions such as "versionInfo," "backwardCompatibleWith," "isDefinedBy," and so on. Typically today, term definitions, their interrelationships, and their associated meta information are not captured for many Web documents, and, if they are captured, are not shared across metadata silos within and among collaborating organizations.

Note that the MDA and Semantic Web communities use the term "metadata" differently. For example, assume that we use OWL to declare FinancialInstitution and Bank as classes and to declare that Bank is a subclass of FinancialInstitution.

[7] [BERNERS-LEE 1999]
[8] [BERNERS-LEE ET AL 2001]
[9] [RDF]
[10] [WEBONT]
[11] [McGUINNESS and van HARMELEN, 2004]]
[12] [OWL GUIDE 2004]

Assume further that we declare CitiBank to be an instance of Bank. In the MDA world, the FinancialInstitution and Bank classes and the subclass relationship between them are metadata that describe CitiBank, which is data. In the Semantic Web community FinancialInstitution, Bank, the subclass relationship, and the declaration of CitiBank as an instance of Bank are "ground facts," whereas metadata about the ground facts includes who authored the facts, when the facts were declared, and so forth. Semantic Web applications reason only about the facts.

Simplified Common Logic (SCL)

RDF and OWL, in and of themselves, do not support full *first-order logic* (FOL) or higher order logics, which are necessary to express certain kinds of facts or constraints. FOL supports *quantification*, which is the ability to express that, for some set of things, "there exists" a member of the set for which a certain condition holds, or that, "for all" the members of some set of things, a certain condition holds. Simplified Common Logic (SCL)[13], which is a nascent successor to the widely used Knowledge Interchange Format (KIF)[14], is a proposed logic standard intended to provide this functionality.

SCL is more expressive than OWL/RDF: there is a standard meaning-preserving translation from OWL/RDF into SCL. OWL/RDF amounts to a collection of packaged forms for particular SCL sentences, producing a more complex syntax but providing advantages for some reasoning engines.

Reasoning over the formal specifications that constitute ontologies can be considerably more challenging when FOL expressions are thrown into the mix. However, the degree of certainty that automated reasoning can achieve need not always be 100%. Raising a red flag that a constraint expressed by an ontology might have been violated is less powerful than being certain about it, but can still be a valuable service.

The syntactic and semantic simplicity of SCL, compared to OWL/RDF, together with its expressive universality, gives it a central role as a core specification of meaning for relating other semantic web languages to one another.[15]

POTENTIAL FOR SYNERGY

Since the modeling of information and knowledge figures prominently in the architecture of both MDA and the Semantic Web, it is important to understand the relationship between the two in order assess the potential for synergy on the one hand and the danger of working at cross-purposes on the other.

Some Overlap

There is some overlap between MDA and the Semantic Web. MOF and UML have some of the same kinds of modeling constructs that the Semantic Web languages have, such as the notion of classes and subclassing. Both also make it possible to model relationships among classes. Although class modeling with RDF and OWL is not identical to class modeling with MOF and UML, it is fair to say that these are overlapping capabilities.

[13] [SCL 2004]
[14] [GENESERETH AND FIKES, 1992]
[15] [LBASE 2003]

Complementary Focuses

Despite some degree of overlap, MOF's focus on automated physical management of metadata is quite different than the Semantic Web's focus on automated reasoning over the resource descriptions. MOF supplies little to support the kind of automated reasoning that Semantic Web technologies support, and the Semantic Web supplies little to automate metadata management.

MDA and the Semantic Web come out of two different communities. MDA reflects the concerns and backgrounds of the community that develops industrial software applications for corporate and governmental IT and for real time and embedded systems. The Semantic Web reflects the concerns of the *knowledge representation* community that is concerned with the problem of how to precisely describe term meanings and how to automatically determine implications of statements that use these terms.[16]

From Research to Industry

Semantic Web technologies are beginning to move from the research phase into industrial applications. Diverse fields, including finance, security, and communications, have needs for precisely defining controlled vocabularies for interoperation and, further, for using inference to detect policy violations or patterns in transactions and transmissions of all sorts.

Companies are also starting to recognize the need to manage metadata that is linked to numerous terminology standards, including generic as well as industry-specific taxonomies or controlled vocabulary specifications, rather than using a single reference. Knowledge representation technologies provide mechanisms for working with multiple terminologies so that applications may specify and recognize relationships between terms.

As ontologies move into industry they need to coexist with industrial metadata. We do not want ontologies to become yet another silo in a fragmented metadata landscape. Since much enterprise tooling is moving toward MOF-based metadata management, a minimal goal would be to make it possible for MOF-based tools to physically manage ontologies using the common MOF mechanisms.

Tapping Existing Metadata

Industry has enormous investments in entity-relationship models, UML models, and relational database models. Such pre-existing sources can serve as starting points for constructing Semantic Web ontologies. It only makes sense to leverage these assets where they exist.

Therefore, we need standardized mappings between these "starter" technologies and the Semantic Web languages, to allow us to mine these valuable sources. Basing such mappings on MOF makes it possible for MOF tools to implement the mappings, thus minimizing the need to hand code the detailed transformations, and helping to integrate the mappings with other MOF-based mappings that will be increasingly important in industry.

[16] For background on knowledge representation, see [BAADER ET AL 2003] and [SOWA 2000]

Scalability of Design-by-Contract™

A number of MDA practitioners emphasize the importance of going beyond the usual bounds of UML modeling when defining the contracts that software components expose via their interfaces[17]. The advanced techniques that they use have some limitations that Semantic Web technology can mitigate.

Quick Introduction to Design by Contract

Consider Figure 1 and Figure 2, which are fragments of a UML model of a banking system's information and available services. The model is at a rather high abstraction level and could be implemented via WSDL or other service-oriented technology. The shaded boxes contain formal declarations of constraints, written in UML's Object Constraint Language (OCL)[18]. The constraint in Figure 1 is an *invariant rule,* which states a condition that must be true for preferred checking accounts. The constraints in Figure 2 are *preconditions* and *postconditions,* also expressed in OCL. The preconditions must be true in order for the service, which transfers funds from a designated checking account to a designated savings account, to behave in a well-defined manner. The postconditions must be true when the service has completed execution.

The use of formal invariant rules, preconditions, and postconditions is a design practice dubbed Design-by-Contract (DBC). DBC is an important technique for improving quality in software engineering and business modeling.

A majority of UML modelers do not use DBC today; in terms of Figure 1 and Figure 2, they model the kinds of things that the unshaded boxes contain, which, in sum, specify the signature of the service interface. In most UML models the constraints show up only in informal text in narrative specification documents that accompany the model diagrams.

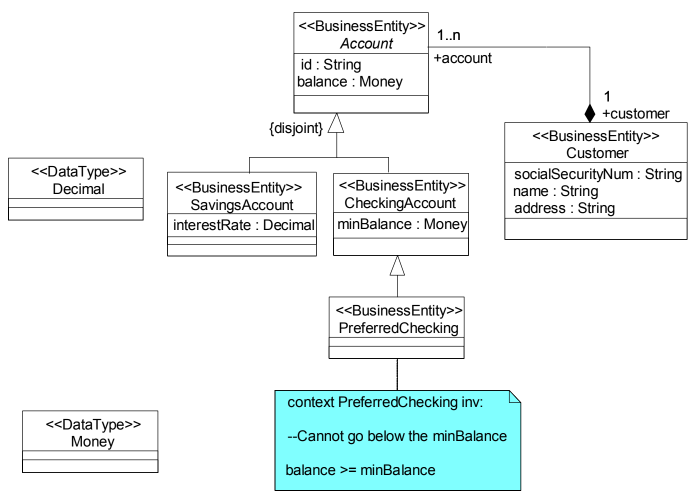

Figure 1: Business Information Model

[17] See [MEYER 1997], [D 'SOUZA AND WILLS 1999], and [FRANKEL 2003]

[18] MDA permits practitioners to practice DBC using languages other than OCL.

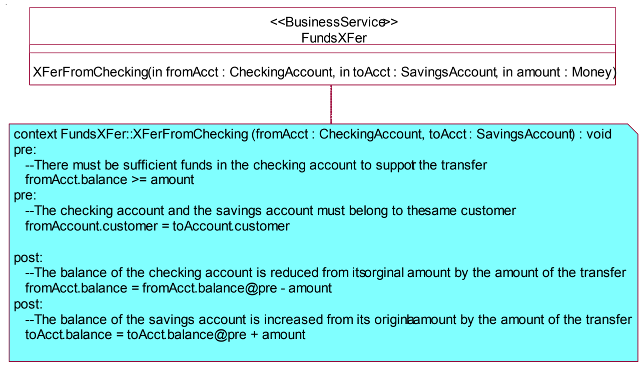

Figure 2: Business Service Model

MDA emphasizes the importance of DBC for two reasons.

1. The increasing pervasiveness of value chains in the economy requires IT to develop service-oriented architectures that manage the links in the chains. Links in value chains consist of services that organizations offer via interfaces. A clear understanding of the contract that obliges the offering organization is critical to rapid and accurate construction of links and to managing the links over time. A formalized contract can serve as the basis for service level agreements (SLAs) that govern relationships across value chains.

2. Formal constraints in models make it possible for model compilers to generate a considerable amount of code that checks whether the constraints are satisfied, or for model interpreters to check constraints dynamically. In value chains, such constraint checking essentially monitors compliance with SLAs. This is especially true when using DBC at a high level of abstraction where implementations details are generally not visible.

Overcoming DBC Limitations

In large, complex systems, diligent practice of DBC results in the proliferation of a large number of invariant rules, preconditions, and postconditions. Syntax checkers for constraint languages catch a certain number of bugs in these rules. However, MDA offers no help in automating the process of determining whether various rules conflict with each other. Conflicting rules generate internally inconsistent implementations.

Semantic Web technology can help with this problem. DBC constraint languages, which support predicate calculus, are significantly less expressive than full blown programming languages such as Java, C#, and so on. This makes it easier to automate reasoning about constraints than it is to reason about a Java or C# program. By using succinct constraint languages to distill the essential, implementation-independent information and behavior of a system, we set the stage for Semantic Web reasoners to ferret out inconsistencies. Of course, arbitrary Java and C# programs may contain inconsistencies, but automating the discovery of logical inconsistencies in such programs is not nearly as tractable.

There is work already in progress in the Semantic Web community to define an approach called *Semantic Web Services,*[19] which seeks to enhance the definition of Web services by going beyond WSDL syntactic definitions to include specifications of preconditions and postconditions for Web service operations, and invariant rules for Web service messages. Direct collaboration between this project and the MDA community has not begun, but there is a strong case for it.

Using UML Notation for Ontology Definition

The W3C has not recommended one graphical notation for use with the Semantic Web languages. Ontologists and knowledge engineers have long recognized the need for a graphical notation, particularly for collaborative development with subject matter experts. In practice, approaches range from formal conceptual graphs or object role modeling, to concept maps, to ad hoc use of desktop drawing tools.

There are a significant number of situations in which there is value in using UML notation and tools for defining ontologies with the Semantic Web languages. Various members of the Semantic Web community have made several proposals for using UML as a presentation syntax for DAML+OIL[20], and, subsequently, for OWL, over the last four years[21]. The value lies in being able to use the widespread knowledge of UML in the software industry, and in leveraging investments that companies have made in UML tools. UML profiling makes it possible to define a dialect of UML that can serve as a graphical notation for the Semantic Web languages.

There are, of course, tradeoffs between using UML notation as opposed to using notations that more precisely target the Semantic Web languages.

Synergy Summary

- *MOF's metadata management facilities and the Semantic Web's reasoning capabilities are separate, complementary concerns*. We want to mine industrial metadata to provide grist for the reasoning mill. MOF alone does not enable us to reason about such metadata; however, MOF can automate a good deal of the process of importing such metadata into Semantic Web ontologies so that automated reasoners can do their work. MOF can also automate the management of Semantic Web ontologies themselves.

- *The ability of Semantic Web reasoners to find discrepancies in invariant rules, preconditions, and postconditions, can add scalability to MDA's use of Design-by-Contract (DBC).* When we combine this kind of reasoning with DBC we significantly increase the potential for DBC to meet its goal of quality improvement.

- *UML profiles can serve as graphical notations for Semantic Web languages.* Using a UML-based graphical notation may not be optimal in all circumstances, but in many cases it has substantial advantages.

---
[19] [SWS]
[20] [CONNOLLY ET AL, 2001]
[21] DAML+OIL is a predecessor of OWL. See [CRANEFIELD AND PURVIS 1999], [KOGUT ET AL 2002], [KENDALL ET AL 2002], [CECCARONI AND KENDALL 2003] for discussion of UML as a presentation syntax for DAML+OIL.

TOWARD A MODEL-DRIVEN SEMANTIC WEB

In order to achieve the goal of using MDA and the Semantic Web together, the OMG issued an RFP that calls for standardizing the following:

- A MOF metamodel for ontology definition
- A UML profile for ontology definition
- A mapping between the UML profile and the MOF metamodel

The RFP mandated that the MOF metamodel cover at least OWL.

Project Participants

The OMG received several initial submissions containing proposed metamodels and profiles. The submitting companies decided to merge their efforts to produce a single, revised submission, which is due to the OMG by 11 October 2004. The official submitting companies are:

- *Distributed Systems Technology Centre (represented by Robert Columb)* : An Australian technology institute that has made many contributions to OMG standards and serves as the Australian office for the W3C
- *Gentleware AG (represented by Marko Boger):* Vendor of a UML tool named Poseidon, which is based on the open source ArgoUML project
- *IBM (represented by Dan Chang):* A key backer of MDA via its Rational and Eclipse tooling
- *Sandpiper Software (represented by Elisa Kendall):* A vendor of ontology tooling, including tools that support using UML to define OWL-based ontologies.

Companies that are unofficially involved with the project are:

- *AT&T Government Solutions (represented by Lewis Hart and Patrick Emery):* An arm of AT&T that has been intimately involved in the Semantic Web
- *David Frankel Consulting (represented by David Frankel)*: A provider of consulting services for enterprise architecture and Model Driven Architecture
- *Institute for Human and Machine Cognition (represented by Pat Hayes):* A research institute of the University of West Florida, among whose personnel is a key author of Semantic Web standards
- *Knowledge Systems Laboratory, Stanford University (represented by Deborah McGuinness)*: Senior personnel of this laboratory have been key authors of Semantic Web standards

Basic Capabilities

Figure 3 captures the technical bridging scenarios that the project intends to support.

Translating Native OWL XML Documents to XMI OWL Documents and Vice Versa

The OWL-XMI Bridge in Figure 3 supports this function.

"Native" OWL XML documents conform to the W3C XML concrete syntax for OWL. XMI OWL documents conform to the XMI-based XML concrete syntax for OWL generated automatically from a MOF metamodel of OWL. The MOF metamodel of OWL is one of the key artifacts that the OMG project is producing. With an OWL-XMI bridge in place, enterprise tooling such as Eclipse can manage OWL ontologies using standard MOF/XMI techniques, and OWL ontologies can be part of the enterprise ecosystem. At the same time, knowledge engineers can use native OWL tooling to define, maintain, and reason over ontologies.

Translating Pre-Existing UML Models to OWL

The OWL-UML Bridge in Figure 3 supports this function.

Pre-existing UML models created with generic UML, rather than with a UML profile for ontology, are valuable assets that we wish to mine by translating them into OWL. Knowledge engineers can then use native OWL tools to refine, extend, and reason over the ontologies.

The project is defining a UML-to-OWL mapping to support this kind of bridge. Certain parts of UML, such as state machines, arguably go beyond ontology definition, and thus the mapping will not cover all of UML.

Using UML Tools to Define OWL Ontologies

The OWL-UML Bridge in Figure 3 supports this function.

A UML profile that the project is developing standardizes the techniques for using UML as a notation for OWL ontology definition. Modelers can use such a profile with generic UML tools.

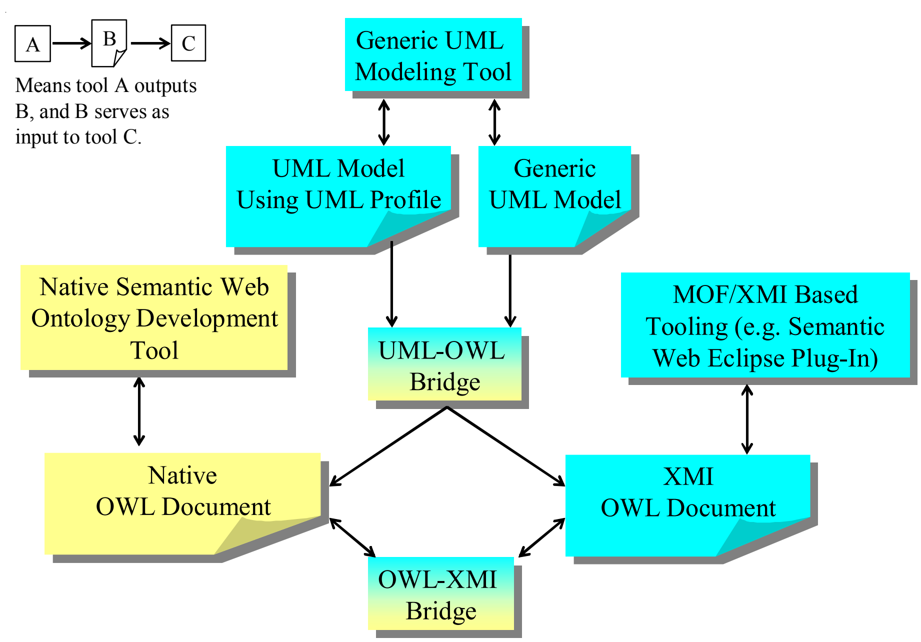

Figure 3: Bridging Between MDA and the Semantic Web

The project is also defining a mapping from the UML profile to OWL, so that add-ins to UML tools can translate UML models based on the profile into native OWL or XMI OWL documents. This mapping is closely related to, but not exactly the same as the generic UML-to-OWL mapping.

Translating Entity-Relationship Models to OWL
The ER-OWL Bridge in Figure 4 supports this function.

As indicated earlier, we wish to mine entity-relationship (ER) models as starting points for ontology development. ER modelers use long-standing ER tools.

The project is defining a MOF ER metamodel and an ER-OWL mapping in order to support this kind of bridge. The source for the mapping is the MOF ER metamodel, and the target is the OWL metamodel.

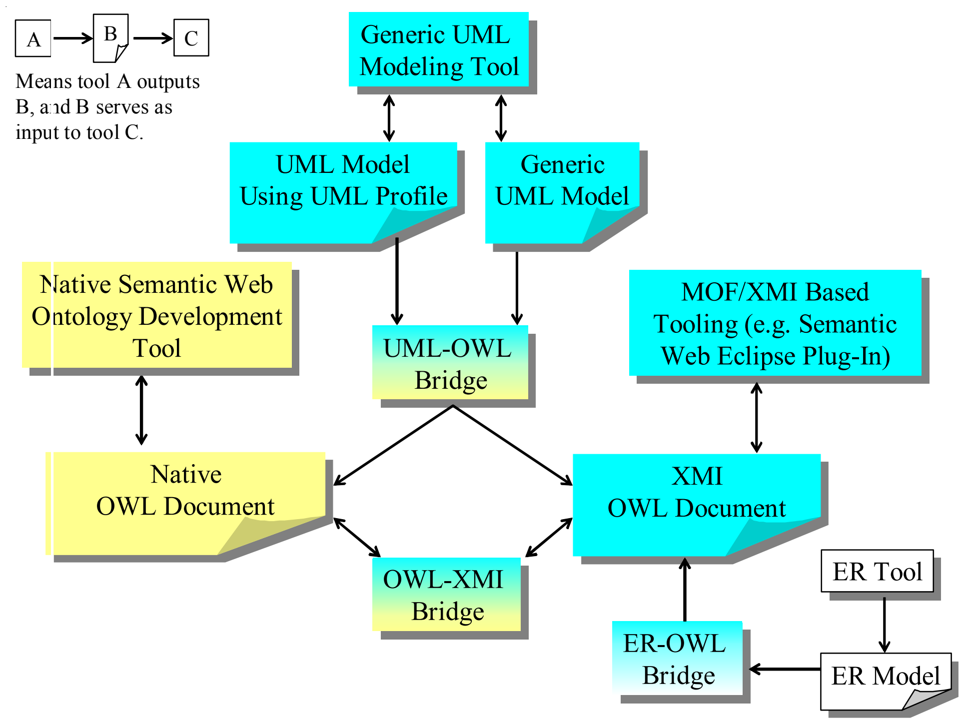

Figure 4: Translating Entity-Relationship Models to OWL

Summary of the Artifacts that the Project is Producing

This section summarizes the artifacts that the OMG project is producing in order to support the bridging capabilities outlined above.

MOF Metamodel of OWL

Although RDF Schema (RDFS) was conceived as an extension of RDF, and OWL as an extension of RDF, in practice it is not feasible to separate RDF and RDFS in order to obtain an RDFS metamodel that is dependent on RDF without the RDF metamodel being dependent on the RDFS metamodel. There are cross dependencies in the W3C specifications of RDF and RDFS that stand in the way.

Therefore, the project is defining one combined RDF/S metamodel and an OWL metamodel that depends on it.

OWL Metamodel – OWL Mapping

This bi-directional mapping defines the relationship between constructs in the MOF metamodel of OWL and the constructs of the OWL language. It is a two way mapping. It is crucial to the bi-directional OWL-XMI Bridge of Figure 3.

UML Profile for OWL

This profile contains a profile for RDF/S and a profile for OWL that depends on it.

UML Profile – to – OWL Mapping

This is a mapping for which the source is the UML Profile for OWL that the project is defining, and the target is the OWL metamodel. It supports translating UML models built via the profile into OWL.

UML – to - OWL Mapping

This is a mapping from the OMG's MOF metamodel of UML, to the MOF metamodel of OWL that the project is defining.[22] It supports translating generic UML models to OWL ontologies.

Simplified Common Logic (SCL) Metamodel

This is a MOF metamodel of SCL. The project positions SCL as a predicate language for enhancing the expressiveness of OWL ontologies. Although SCL's authors intend to use SCL this way, SCL does not depend on OWL. Therefore, the SCL metamodel does not depend on the OWL metamodel.

Entity Relationship (ER) Metamodel

This is a MOF metamodel of ER.

ER-OWL Mapping

This is a mapping from the ER metamodel to the OWL metamodel.

XML Topic Maps

XML Topic Maps are another ontology definition language[23] that may be important to the Semantic Web. If the project team has sufficient bandwidth, it will define a MOF metamodel for Topic Maps, a UML Profile for Topic Maps, and corresponding mappings.

Core Metamodel

In defining the mappings listed above, the project team wishes to avoid simply defining point-to-point mappings that lead to classic N^2 mapping explosion. Thus, the team is exploring the feasibility of defining a Core metamodel that contains elements common to all of the metamodels. Most likely, this Core will be based on *description logic*, a constrained form of knowledge representation that is particularly amenable to supporting automated reasoning[24]. If the team decides to proceed with this approach, it will define a series of bi-directional mappings, where each mapping is a mapping between the Core metamodel on the one

[22] The project team may decide to define a separate UML-RDF mapping to support translating UML models directly into RDF without the OWL abstractions.
[23] [TOPIC MAPS]
[24] See [BAADER ET AL 2003] for a comprehensive introduction to description logic.

hand and either the OWL, UML, ER, or Topic Maps metamodel on the other (see Figure 5).

Consider a mapping from metamodel A to metamodel B, where neither metamodel is the core metamodel. Using this core approach, we would construct the mapping from the A-->Core mapping and the Core-->B mapping.

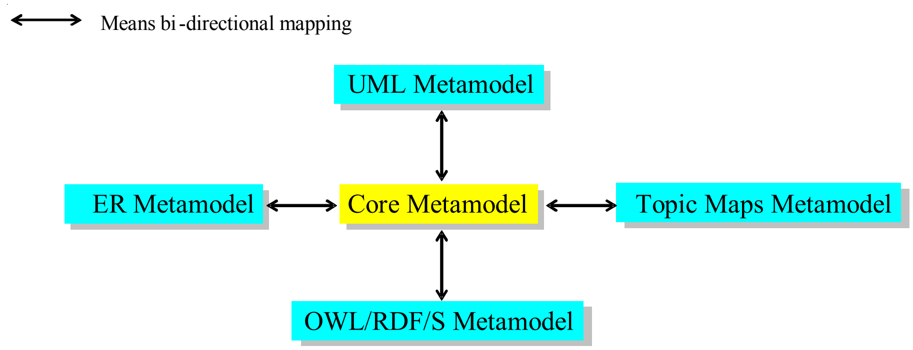

Figure 5: The Core Metamodel

The team is exploring whether to use the MOF Query View Transformation (QVT) standard to define the mappings. The QVT standard is still a work in progress, but will make it possible for QVT-based transformation engines to use standardized techniques to preserve mapping information after translating a model expressed in one language to a model expressed in another language. For example, if such an engine translates a particular UML model to an OWL ontology by applying the UMLàCore and CoreàOWL mapping, it can maintain tracings that make it possible to determine that a particular element in the resultant OWL ontology traces back to a particular element in the UML model. Such traceability obviates the need for the Core to support all of the semantics of all of the other metamodels in order to avoid losing information when executing transformations.

INDUSTRIAL APPLICATIONS

Model-Driven Semantic Web technology has the potential to make an impact in a number of industrial domains. This section provides a brief overview of some industrial applications.

Finance

Here we supply two of many possible uses in the finance industry, which manages much complex information and rules.

Portfolios of Capital Market Contracts

Consider a portfolio of holdings that consists of capital market contracts such as options, swaps, and complex derivatives thereof. Knowledge representation technology offers the possibility of defining formal ontologies of such contracts that capture term definitions, their interrelationships, and rules, and can help monitor and maintain the internal consistency of such descriptions.

Furthermore, assessing the value of positions and quantifying the risk that such positions hold depends heavily on additional vocabularies and rules, as does

[25] See [LONGVIEW] for an example of a capital markets software company that is augmenting traditional component-based development with semantic ontology modeling.

breaking down general analyses of value and risk into fine-grained views that highlight the contribution of specific factors to changes in value and risk. The greater the extent to which we can capture the information vocabularies and contract rules as formal ontologies, the greater our capacity to automate assessments of risk and value.[25]

At the same time, such ontologies must co-exist with metadata in corporate finance systems, particularly entity-relationship models and relational database models. As indicated earlier, we should like to be able to deploy these ontologies in a manner that avoids creating another silo in the enterprise, and an MDA-based approach to using ontologies offers important value in this respect.

Furthermore, a system of automation that uses formal ontologies will have to map the ontologies to back end data stores. We would prefer to define these mappings as mappings of source and target MOF metamodels and to generate the transformation code from the mappings. Using this model-driven approach, transformations can evolve gracefully, as ontologies and back end data stores evolve, with less dependence on manually maintaining lower level transformation code.

Security

The finance industry is well aware that traditional authentication and authorization schemes, while important, are not sufficient to guard against fraud and other forms of security violations. More sophisticated rule-based systems protect against identity theft, money laundering, capital flows that threaten national security, and so on. The formally grounded mathematical techniques wherein lies the expertise of the knowledge representation community has a lot to offer in improving such detection schemes.

We also know that links in value chains—that is, the points where systems from different organizations interact with each other—are key security vulnerability points. As value chain business continues to spread, an organization is likely to have to maintain many value chain links at which the organization plays some role. The greater the extent to which we can automate the implementation and monitoring of these links, the easier it becomes to control them. When the implementation of each link is buried deep in source code rather than being manifest in a consistent, automatically replicated link model, the potential is greater for back doors—whether created deliberately or inadvertently—to escape detection.

A key to automating the links is the ability to model them formally. MDA and the Semantic Web Services project both offer means to do so. And, here again, formal ontologies cannot exist in a vacuum—they must co-exist and interact with other information in the enterprise, which again highlights the importance of the MDA side of the equation.

Policy and Compliance Management

Several ongoing applied research programs are using knowledge representation technologies to declaratively represent and reason over policy and compliance knowledge. While near term demonstrations can show dramatic improvements over current capabilities using knowledge representation technologies alone, the marriage of MDA and Semantic Web capabilities can support the scalability, maintainability, and robustness requirements of long term deployment and system evolution.

For example, UltraLog is a Defense Advanced Research Projects Agency (DARPA) sponsored research project focused on creating survivable large-scale distributed agent systems capable of operating effectively in very chaotic environments.[26] The objective of the project is to create a comprehensive capability that will enable a massive scale, trusted, distributed agent infrastructure for operational logistics to survive under the most extreme circumstances. One component of that infrastructure, called the KAoS (Knowledgeable Agent-oriented System) Policy and Domain Services, provides for specifying, managing, resolving conflicts among, and enforcing policies within, specific contexts established by complex organizational structures, where the policies and organizational structures are represented as OWL ontologies.[27]

The research uncovered a number of limitations in the tools and technologies that were used, for example, for ontology creation and management, including lack of support for component-based ontology development, inability to link multiple ontology components to one another, inability to ensure transitive closure across imported ontologies and related knowledge bases, and scalability problems. Only some of these issues can be categorized strictly as tool limitations; an integrated MDA approach can not only limit their impact, but also in some cases eliminate them. Furthermore, integration of MDA and Semantic Web Services technologies could provide graphical visualization and automation capabilities supporting the interaction semantics that were found to be a critical component of the KAoS architecture.

One interesting result produced by the KAoS development team was in the use of ontologies to enable dynamic configuration of software components for particular installations. The roles played by the various components (such as enforcers, classifiers, and policy editors) were related to concepts defined in specialized ontologies. These relationships—which constitute additional metadata for the system—were registered in software factories and were used to create Java components on demand. Such an approach, coupled with MDA metadata management technologies, could enable a powerful, next generation software development framework.

MICROSOFT AND THE MODEL-DRIVEN SEMANTIC WEB

Microsoft is working on its own model-driven software technology.[28] Much of what we have stated about the relationship between MDA and the Semantic Web could possibly apply to the relationship between Microsoft's model-based technology and the Semantic Web. It is too early to assess the relationship between Microsoft's modeling direction and MDA, but this bears watching.

---

[26] [ULTRALOG]
[27] [USZOK ET AL 2003]
[28] See [COOK 2004] and [GREENFIELD AND SHORT 2004]. Steve Cook's MDA Journal article was one of the first public statements by Microsoft about its modeling direction. The Greenfield et al book is an important and substantial piece of work, put together by members of the architecture team of which Steve Cook is a part, and Steve made important contributions to it. As of this writing the book had not yet been released, but it should be available soon.

CONCLUSION

We have just begun to explore the potential for MDA and the Semantic Web to strengthen each other.[29] The software industry has started the slow, gradual process of introducing MDA and other model-driven approaches to software, and the Semantic Web also is a technology that industry can assimilate only gradually. We thus have some lead time to align these two communities so that we can offer industry a whole that is more than the sum of its parts.

This is not to advocate delaying MDA and Semantic Web projects pending a total alignment. Furthermore, alignment may never be 100% complete in practice. However, the sooner and the more we can achieve such alignment, the less we will have to undo legacy incompatibilities when we seek to combine the technologies.

References

[BAADER ET AL 2003] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter Patel-Schneider, *The Description Logic Handbook: Theory, Implementation, and Applications*, Cambridge University Press, 2003, ISBN 0521781760.

[BERNERS-LEE 1999] Tim Berners-Lee, *Weaving the Web,* Harper San Francisco, 1999, ISBN 0062515861.

[BERNERS-LEE ET AL 2001] – Tim Berners-Lee, James Hendler, Ora Lassila, *The Semantic Web*, Scientific American, May 2001.

[CECCARONI AND KENDALL 2003] Luigi Ceccaroni and Elisa Kendall, "A Graphical Environment for Ontology Development", *Proceedings, The Second International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS 2003)*, Poster Session, July 14-18, 2003, Melbourne, Australia, The ACM Digital Library.

[CMI] MOF 2.0 IDL, OMG document ad/04-01-16.

[CONNOLLY ET AL, 2001] Dan Connolly, Frank van Harmelen, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. DAML+OIL (March 2001) Reference Description. World Wide Web Committee (W3C) Note 18 December 2001. http://www.w3.org/TR/daml+oil-reference

[COOK 2004] Steve Cook, *Domain-Specific Modeling and Model Driven Archit*ecture, MDA Journal, January 2004, Business Process Trends.

[CRANEFIELD AND PURVIS 1999] Stephen Cranefield and Martin Purvis, "UML as an Ontology Modelling Language", in *Proceedings of the Workshop on Intelligent Information Integration, 16th International Joint Conference on Artificial Intelligence (IJCAI 99)*, 1999.

[DACONTA ET AL 2003] Michael C. Daconta, Leo J. Orbst, and Kevin T. Smith, *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*, Wiley Publishing Inc., 2003, ISBN 0471432571.

[D'SOUZA AND WILLS 1999] Desmond D'Souza and Alan Cameron Wills, *Objects, Components, and Frameworks with UML: The Catalysis Approach*, Addison Wesley, 1999, ISBN 0201310120.

[ECLIPSE] http://www.eclipse.org

[FRANKEL 2003] David S. Frankel, *Model Driven Architecture: Applying MDA to Enterprise Computing*, John Wiley & Sons, 2003, ISBN 0471319201.

[29] For another perspective on the potential for the Semantic Web in the enterprise, see [LO GIUDICE AND GUTTMAN 2003]

[GENESERETH AND FIKES 1992] Michael Genesereth & Richard Fikes, "Knowledge Interchange Format, Version 3.0 Reference Manual", KSL Report KSL-92-86, Knowledge Systems Laboratory, Stanford University, June 1992.

[GREENFIELD AND SHORT 2004] Jack Greenfield and Keith Short, with Steve Cook and Stuart Kent, *Software Factories: Assembling Applications with Patterns, Models, Frameworks, and Tools*, John Wiley & Sons, 2004, ISBN 0471202843.

[JMI] *Java Metadata Interface Specification*, Java Specification Request 40, Java Community Process, http://jcp.org/jsr/detail/40.jsp

[KENDALL ET AL 2002] Elisa F. Kendall, Mark E. Dutra, and Deborah L. McGuinness, "Towards a Commercial Ontology Development Environment", In *Proceedings of the International Semantic Web Conference – Late Breaking Topics*, Sardinia, Italy, June 9-12, 2002. http://www.ksl.stanford.edu/people/dlm/papers/iswc-sandpiper-abstract.html

[KOGUT ET AL 2002] Paul Kogut, Stephen Cranefield, Lewis Hart, Mark Dutra, Kenneth Baclawski, Mieczyslaw Kokar, and Jeffrey Smith, " UML for Ontology Development", In *Knowledge Engineering Journal*, Volume 17, Issue 1, March 2002, pp. 61.64.

[LBASE 2003] R.V. Guha and Patrick Hayes, "LBase: Semantics for Languages of the Semantic Web," http://www.w3.org/TR/lbase/

[LO GIUDICE AND GUTTMAN 2003] Diego Lo Giudice and Michael Guttman, "Where Enterprise Architecture Meets the Semantic Web," Cutter Enterprise Architecture Advisory Service, Executive Report Vol. 6., No. 8.

[LONGVIEW] Longview International Inc., www.lvi.com

[MCGUINNESS AND van HARMELEN 2004] Deborah L. McGuinness and Frank van Harmelen, editors, OWL Web Ontology Language Overview, W3C Recommendation 10 February 2004, http://www.w3.org/TR/2004/REC-owl-features-20040210/.

[MEYER 1997] Meyer, Bertrand, *Object-Oriented Software Construction, Second Edition*, Prentice-Hall, 1997, ISBN 0136291554.

[OWL GUIDE 2004] Michael K. Smith, Chris Welty, Deborah L. McGuinness, *OWL Web Ontology Language Guide*, W3C Recommendation 10 February 2004, http://www.w3.org/TR/2004/REC-owl-guide-20040210/

[RDF] Resource Description Framework, http://www.w3.org/RDF

[SCL 2004] Patrick Hayes, *Abstract Syntax and Semantics for SCL*, http://cl.tamu.edu/docs/scl/scl-latest.html

[SOWA 2000] John F. Sowa, *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks/Cole, 2000, ISBN 0534949657.

[SWS] Semantic Web Services, http://www.daml.org/services/

[TOPIC MAPS] XML Topic Maps (XTM), http://www.topicmaps.org

[ULTRALOG] A DARPA Program on Logistics Information System Survivability. See http://www.ultralog.net/.

[USZOK ET AL 2003] Andrzej Uszok, Jeffrey M. Bradshaw, Patrick Hayes, Renia Jeffers, Matt Johnson, Shriniwas Kulkarni, Maggie R. Breedy, James Lott, and L. Bunch (2003). DAML reality check: A case study of KAoS domain and policy services. Submitted to the International Semantic Web Conference (ISWC 03). Sanibel Island, Florida.

[VON SCHWEBER 2001] Erick Von Schweber, "Towards a Model Driven Semantic Grid?," presented to the OMG Software Services Grid Workshop and the Global Grid Forum, July, 2001, http://www.synsyta.com/readings/docs/infomaniacs/software-services-grid-workshop-update-09-2002.ppt

[W3C SEMANTIC WEB ACTIVITY] http://www.w3.org/2001/sw/
[WEBONT] W3C Web-Ontology (WebOnt) Working Group, http://www.w3.org/2001/sw/WebOnt
[XMI] XML Metadata Interchange, Version 2.0, OMG document formal/03-05-02.

AUTHOR BIOS

David Frankel

www.DavidFrankelConsulting.com/bio.htm

David Frankel's career in the software industry spans 25 years, during which he has had experience in all phases of software development, including requirements gathering, writing of specifications, formal design, coding, testing, internal and user documentation, design and teaching of training courses, deployment, and long-term maintenance. He specializes in the architecture of distributed enterprise computing systems. He is the author of many published articles and sole author of the book *Model-Driven Architecture: Applying MDA to Enterprise Computing*, published by John Wiley & Sons in January 2003.

He served several terms as an elected member of OMG Architecture Board, and was intimately involved in the launch of MDA. He is the co-author of several industry standards, including COM-CORBA Interworking, the UML™ Profile for CORBA®, and the UML Profile for EJB™. He is the owner of David Frankel Consulting. He has a Bachelor's degree in mathematics and a Masters degree in social work, both from the University of Illinois.

Patrick Hayes

http://www.ihmc.us/users/phayes

Pat Hayes received a BA in mathematics from Cambridge University and a PhD in Artificial Intelligence from Edinburgh. He has held academic positions in computer science at the University of Essex (England), in philosophy at the University of Illinois and as the Luce Professor of cognitive science at the University of Rochester. He has been a visiting scholar at Universite de Geneve and the Center for Advanced Study in the Behavioral Studies at Stanford, and has directed applied AI research at Xerox-PARC, SRI and Schlumberger, Inc. At various times, Pat has been secretary of Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB), chairman and trustee of the International Joint Conference on Artificial Intelligence (IJCAI), associate editor of Artificial Intelligence, a governor of the Cognitive Science Society and president of AAAI.

Pat's current research interests include knowledge representation and automatic reasoning, especially the representation of space and time; the semantic web; ontology design; and the philosophical foundations of AI and computer science. He also restores antique mechanical clocks, remodels old houses, draws portraits and enjoys arguing with anyone about almost anything. Pat is a charter Fellow of the American Association for Artificial Intelligence (AAAI) and of the Cognitive Science Society, and has professional competence in domestic plumbing, carpentry and electrical work.

Elisa Kendall

http://www.sandsoft.com/management.html

Elisa Kendall is Chairman & CEO of Sandpiper Software, which she founded in 1995. She has over 25 years professional experience in the design, development, and deployment of enterprise-scale information management systems for communications, high technology, and aerospace applications. She is the principal architect of Sandpiper's UML-based knowledge representation, ontology analysis, and reasoning architecture, and has developed rigorous methodologies for domain assessment and ontology design and development. Ms. Kendall currently leads the policy awareness segment of Raytheon's DARPA/XG program for next generation communications, is a key contributor to standards efforts such as the OMG's Ontology Definition Metamodel (ODM), and provides training and services to several of the company's key customers.

From August 1992 to June 1995, she served as product development manager for Aspect Development, a provider of component and supplier management systems acquired by i2 Technologies. Prior to joining Aspect, Ms. Kendall served in a variety of roles at Lockheed Martin Missiles & Space. She was lead software engineer for the design, development, deployment, and support of several ground-based signal processing and data management systems. She holds an MA in Linguistics from Stanford University and a BS in Mathematics and Computer Science from the University of California at Los Angeles.

Deborah McGuinness

http://www.ksl.stanford.edu/people/dlm/publicity.html

Deborah McGuinness is the associate director and senior research scientist of the Knowledge Systems Laboratory at Stanford University. She has been working in knowledge representation and reasoning environments for ontology creation and maintenance for over 20 years. She has built and deployed numerous ontology environments and ontology applications, including some that have been in continuous use for over a decade at AT&T and Lucent. She is the co-editor of the W3C Recommendation Ontology Markup Language (OWL) and co-author of the predecessor languages: the DARPA agent markup language (DAML+OIL), OIL, and CLASSIC. She leads the Stanford Explanation and Ontology Evolution Environment efforts. She has published over 100 papers and has authored granted patents in knowledge based systems, ontology environments, configuration, and search technology.

Deborah's consulting business helps companies plan, develop, deploy, and maintain semantic web applications. Some areas of recent work include: ontology environments, search, eCommerce, eHealth, configuration, and supply chain management. She is on the advisory board for Network Inference, Radar Software, Sandpiper Software, and Buildfolio, and recently advised Applied Semantics and Guru Worldwide prior to their acquisitions. Deborah is program chair for the 2004 American Association for Artificial Intelligence conference, and she is on the steering board for some other academic organizations, including the International Organization for Description Logics (DL), the Semantic Web Science Foundation, Ontology.org, the International Organization for Knowledge Representation and Reasoning (KR inc.), and the International conference on Conceptual Structures (ICCS). Deborah received her Bachelors degree in math and computer science from Duke University, her Masters degree in computer science from Berkeley, and her PhD from Rutgers University.

"CORBA," "MDA," "Model Driven Architecture," "UML," and "XMI" are registered trademarks of the Object Management Group.
"Design by Contract" is a trademark of Interactive Software Engineering.
"Java" is a trademark of Sun Microsystems.
"MOF" is a trademark of the Object Management Group.