Perspectives on Applying Semantic Web Technologies in Military Domains

Mary Pulvermacher, Leo Obrst, Salim Semy, and Suzette Stoutenburg

1. Purpose. This paper summarizes our current thoughts, opinions, and lessons learned regarding Semantic Web technologies. The recent advances in these technologies, combined with their rapid pace of change, lead us to believe that there are compelling potential benefits for our Military customers. Furthermore, these technologies are starting to be applied today both in the Government and commercially. This paper is intended to serve as an executive overview for MITRE and U.S. Government decision makers.

Semantic Web Technologies Executive Overview

2. Our Background. Our opinions have been formed based upon a foundation of formal education as well as direct experience using Semantic Web technologies, especially over the last two years through MITRE-sponsored research.

> **"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."**
> *Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, May 2001*

3. What is the Semantic Web? The Semantic Web is defined as "an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." The Semantic Web extends the current World Wide Web incrementally. The current Web is designed to present information to people. The Semantic Web extends the current Web to make web information meaningful to machines (software) by attaching some well defined meaning (i.e., semantics) to the information. These semantics are made explicit through ontologies that are then exploited by software applications.

Ontologies are the next step in increasingly strong knowledge capture. An ontology defines terms and concepts and their relationships to one another precisely enough to support machine interpretation. A concept may be thought of as a resource identified by a Uniform Resource Identifier (URI). Resources may either exist on the Web (e.g., a document that may be retrieved) or be represented on the Web (e.g., a person). With their adoption by the Web community, ontologies founded upon Web technologies have a chance at ubiquity. A useful definition of ontology may be found in [OWL03, pp. 3-4] and [DOS03, pp. 166-167 and 185-188]. A discussion of what ontology means from the perspective of information technology appears in [GG95].

Web ontology languages are founded on a language called Resource Description Framework (RDF)[1] and its subsequent extension called RDF Schema (RDFS), together referred to as RDF/S. RDF/S represents resources as sets of triples, where each triple consists of either (Resource, Property, Resource) or (Resource, Property, PropertyValue). RDFS is a simple ontology language; RDF triples are instances of classes and properties defined by RDFS. These triples collectively constitute a graph as shown in the simplified vehicle ontology in Figure 1. The Web Ontology Language (OWL)[2] is a semantic extension of RDF/S, providing more expressive power. On 10 February 2004, the W3C announced that RDF and OWL were W3C Recommendations[3], effectively making them the standard Web vocabulary for capturing semantics.

OWL is to semantics as XML is to syntax/structure

[1] http://www.w3.org/RDF/
[2] http://www.w3.org/2004/OWL/
[3] http://www.w3.org/2004/01/sws-pressrelease

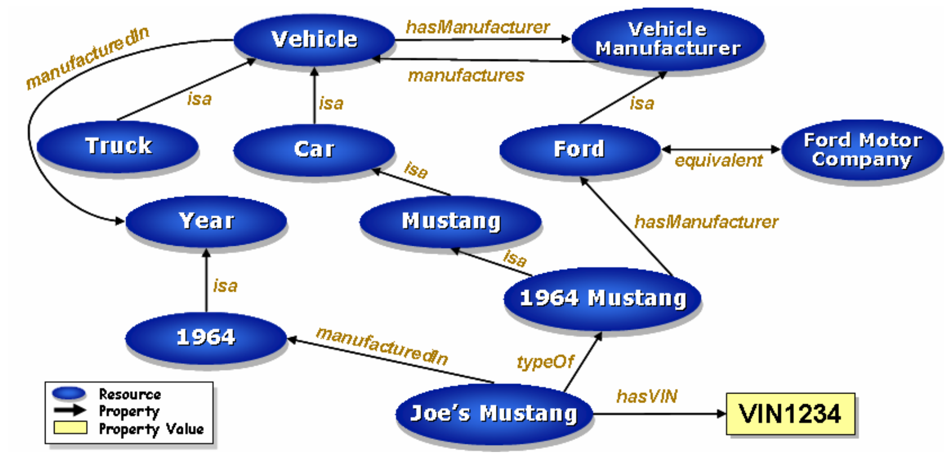

Figure 1. Sample Ontology

Ontologies provide a mechanism for machines to perform simple to complex inferencing by combining facts together to form new facts or conclusions. For example, given that *Mary* is the *spouseOf Jim* and *spouseOf* is a symmetric property, an inference engine can conclude that *Jim* is the *spouseOf Mary* without having that fact explicitly asserted. OWL provides the semantic expressiveness that enables machines to inference on the ontology or on instance data mapped to the ontology, sometimes called the "knowledge base". Inferencing may be based upon things like relations between concepts (equivalent, disjoint, etc.), property characteristics (inverse, transitive, symmetric, etc.) or cardinality constraints (e.g., birthmother has exactly one value). For example, the 'subclass' relation is a transitive property and so inheritance of properties from superclasses to subclasses and to the instances of those classes are supported. Constraints and the capability to combine facts to make inferences allow machines to solve tedious problems, combine facts to discover new information, and help prevent certain misunderstandings. Further, inferencing across ontologies offers the opportunity to achieve a "network effect" across the Web.

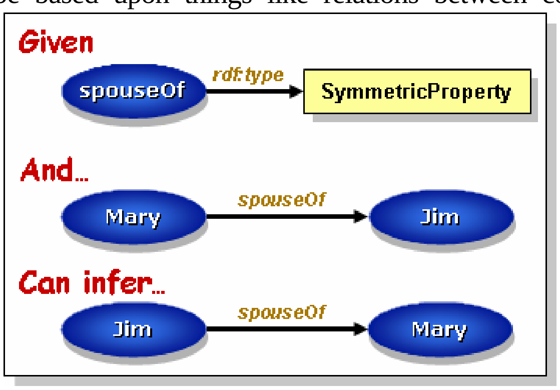

Figure 2. Inferencing Example

4. Why is the Semantic Web Important to Military Customers? Why is the Semantic Web and its associated technologies important to our military customers? First, the associated technologies can be applied to a wide range of today's enterprise challenges, such as Interoperability and Enterprise Integration, even if realization of a ubiquitous Semantic Web is years away. Also, the potential exists for even greater capabilities in the future. Semantic Web technologies enable the ability to specify the meaning of concepts in a standard way across a network. The standard allows for universal interpretation of concepts by machines, thus paving the way for enhanced machine-to-machine (M2M) interaction for enhanced mission effectiveness. For example, easier, quick-to-create M2M interaction across a set of Coalition partners would enable more rapid situational awareness, thereby enhancing the Warfighers' ability to respond and succeed on the battlefield.

Semantic technology for Interoperability and Enterprise Integration

5. Our Perspectives on Semantic Web Maturity. While much work remains, the Semantic Web is maturing rapidly. Our view is supported by our direct experiences, as well as our observations regarding advances being made on international standards, emerging tools, and the proliferation of interest across research, government, and industry domains. Current advances are building a foundation for a ubiquitous Semantic Web and the future looks very promising.

The Semantic Web is today where XML Schema was 5 years ago

We see a number of advances being made to support standardization in multiple facets of the Semantic Web. As mentioned previously, in February 2004, the Web Ontology Language (OWL) became a World Wide Web Consortium (W3C) Recommendation. This has led to a growing movement by researchers, government, and commercial industry to align their efforts to support these standards. More recently, initiatives have begun to define standards in other areas including standard rule and query languages.

We also see a growing interest in developing reusable ontologies to facilitate interoperability. This includes still controversial initiatives to develop a Standard Upper Ontology (SUO) that includes high-level, abstract, domain-independent concepts. This also includes efforts to develop ontologies of commonly used concepts such as time and space, often referred to as a utility ontology, as well as super domain ontologies that contain concepts that span sub-domains. Using common mid-level and upper ontologies is intended to ease the process of integrating or mapping domain ontologies. Figure 3 shows a notional categorization of ontologies. Finally, accessibility and reusability of ontologies are further supported through emerging ontology repositories.

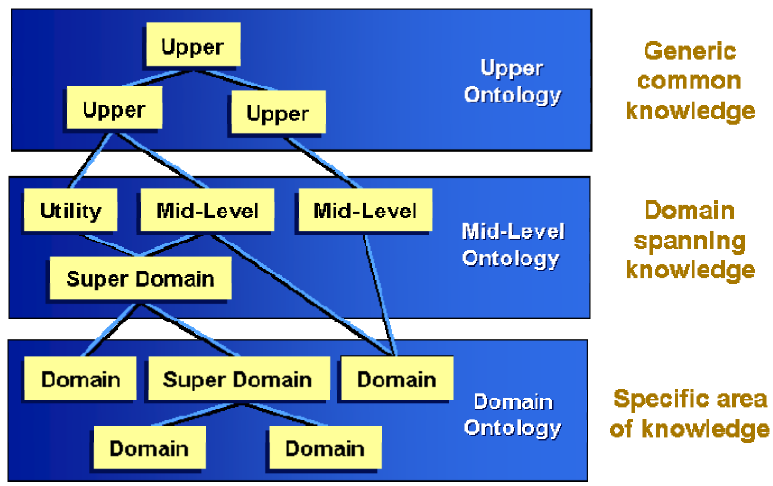

Figure 3. Ontology Categories

The United States Government has been investing heavily in Semantic Web technology over the past decade. Early work by DARPA and the DARPA Agent Markup Language Program (DAML) fueled momentum in this area. The government has also been one of the early adopters of this technology, as illustrated by projects such as the Intelligence Community Metadata Working Group,[4] E-Gov,[5] and various National Security Agency (NSA) projects. NASA recently won the Semantic Web Challenge award at the International Semantic Web Conference 2004 for their work in applying semantic technology to integrate the findings of their disparate scientific programs.[6]

Industry has begun to embrace these technologies. A growing number of vendors are commercializing research tools, developing tools for ontology construction and management, supporting international standards, and applying these standards and tools to practical problems. This movement, in part, has led to an increased awareness and application of this technology in the commercial sector. Examples of early adoption include: Sun's in-house digital asset management system,[7] Adobe's RDF based Extensible

[4] http://www.xml.saic.com/icml/ic_mwg/introduction.asp
[5] http://www.whitehouse.gov/omb/egov/
[6] http://challenge.semanticweb.org
[7] http://sun.systemnews.com/articles/68/4/pr-partner/11092

Metadata Platform (XMP)[8] and Google's planned search engine enhancements through acquisition of Applied Semantics.[9] On the other hand, adoption of these technologies is uneven. To date, mainstream application development, database and middleware vendors, as well as the database research community have been slow to embrace the Semantic Web.

With this rising momentum across government, academia and industry, there has also been a growing investment in the Semantic Web by MITRE. MITRE has been involved in the previously mentioned DOD/Intelligence Community initiatives. MITRE has also supported a number of internal research efforts, supporting both the development of Semantic Web technical foundations as well as working with sponsors to apply this technology. More recently, a corporate Semantic Web Innovation Grant team was formed within MITRE to research the landscape of the Semantic Web and determine it's relevance to MITRE's work program. One of the results of this effort was a recommendation, which was accepted by MITRE's Engineering Council, to make the Semantic Web a MITRE Technology Program thrust. In FY05, we see continued MITRE R&D efforts in areas such as standard rule languages, semantic enterprise integration, and applications of Semantic Web technologies.

This increased awareness of the Semantic Web has led to the evolution of conferences beyond the research community to include heavy participation from industry and government. Conferences of particular dominance in this area include: International Semantic Web Conference,[10] Semantic Web track in the International World Wide Web Conference,[11] and the Formal Ontology in Information Systems Conference.[12] We also see the development of interest groups focused in areas such as Semantic Web Services, data access,[13] best practices and deployment,[14] and a Semantic Interoperability Community of Practice (SICoP).[15]

While there is clear indication that the Semantic Web is maturing rapidly, we believe further maturation is necessary. Our experiences with using vendor and research tools have shown that the tools need to mature further to more completely realize the Semantic Web and that our direct experiences can facilitate this maturation. We also believe standard Semantic Web rules and query languages are needed; MITRE is already starting to participate in these fronts. As these technologies are applied to mission critical, secure environments, issues of trust, privacy and security must be addressed. Because these topics are vital to the government, the government and MITRE should consider increasing participation in these areas and promote their importance to academia and industry. Some specific issues that should be considered are laid out in [SLP03], [PSS04], and [SPO04].

Government and MITRE should invest in Semantic Web technologies for trust, privacy and security

6. Our Perspectives on Where We Believe We'll See Near-Term Value. While the Semantic Web vision is that the current World Wide Web evolves to be a ubiquitous web of semantically tagged data and applications that achieves a similar "network effect", more humble advances will be realized in the near-term. Many of these advances will be in domain-specific applications that use ontologies as conceptual models of their domain.

---

[8] http://www.adobe.com/products/xmp/main.html
[9] http://www.google.com/press/pressrel/applied.html
[10] http://iswc.semanticweb.org/
[11] http://www.iw3c2.org/
[12] http://fois2004.di.unito.it/
[13] http://www.w3.org/2001/sw/DataAccess/
[14] http://www.w3.org/2001/sw/BestPractices/
[15] http://web-services.gov/

Semantic Web Services[16] is an area where we predict significant near-term advances will be realized. Web service use is now a mainstream practice. Describing web services with a standard ontology can help with service discovery and service orchestration. Also, interoperability can be fostered through the use of semantically enabled web services. A human user of a semantically aware web service will likely be unable to ascertain that Semantic Web technologies are being used. The fact that ontologies are being used to discover a needed service or provide context for the data entered will be transparent to a user.

Semantically aware searches and web crawlers are other applications where we envision strong near-term impact. In fact, research prototypes of web crawlers[17] and simplified taxonomies[18] such as Yahoo! directories[19] are used on the Web today. Ontologies provide the potential to transform search engine technology by offering context and semantics to guide the user through query refinement. For example, if one searched on "*mustang*", the engine could offer the user a number of contexts for mustang to narrow the search, such as "*car*", "*horse*", or "*military operation*".

Finally, Semantic Web technologies can support semantic interoperability in the near-term. Semantic Web formalisms (e.g., ontologies, linkages between ontologies, and linkages from ontologies to data sources and services) are well suited to making knowledge captured once, reusable for other purposes. Recent MITRE research (e.g., [PSS04], [SP04], and [SKS04]) supports this view.

For all these applications, it is not clear where the return on investment (ROI) is likely to be greatest. As applications in these and other areas emerge, correspondingly more evidence will become available to support ROI studies.

7. Our Perspectives on Where We Believe We'll See Long-Term Value. As advances are made in Semantic Web technologies, we anticipate their use will be applied to help solve some of the toughest data semantics problems that remain. This includes advances in natural language processing, intelligent agent interaction, enterprise interoperability and other advanced applications such as data mining. We also anticipate that the ability to capture the semantics of enterprise architectural models offers the hope for a more seamless implementation approach. Ideally, one could capture sufficient semantics in operational and system architectures such that applications could be automatically driven by these models. Use of standard Web-based technologies will allow long-term advances to be achieved incrementally by building upon earlier semantic knowledge capture successes.

The Semantic Web – Another Web "network effect"

8. Our Advice on What to Do in the Near-Term. Semantic Web technologies are being used today. However, we believe they won't achieve widespread acceptance until tools mature and some of the enterprise issues are resolved. Furthermore, developing ontologies can require a substantial up-front investment. So what should we do today? Ideally we would like to position ourselves and our customers to realize the power of these technologies while mitigating the impact on legacy systems as the tools mature. Here are our suggestions on a transition strategy.

Monitor innovations in Semantic Web tools and standards. Thousands of people are applying their talents to advancing the art and practice of these technologies and hence, they are maturing rapidly. Stay abreast of these advances to make informed decisions on when they should be applied in your domain.

---

[16] See http://www.daml.org/services/owl-s/ and http://www.wsmo.org/
[17] See SWOOGLE (http://swoogle.umbc.edu/) for example.
[18] A taxonomy may be thought of as a simple form of an ontology, one in which the relationships are strictly hierarchical.
[19] http://dir.yahoo.com/

Grow talent. Given the Semantic Web momentum and the potential offered by these technologies, we believe it is vital to invest in training on Semantic Web technologies. This includes developing data modeling and ontological engineering expertise. Ontologies are powerful but their development and use are non-trivial.

Capture semantics in an ontology. Consider capturing data and application semantics in an ontology. Data and application semantics exist today. Capture them explicitly in a standard, machine interpretable way. Develop these ontologies with reuse in mind and consider registering the ontologies to make them available to others.

Model data based on the real-world. Where possible, model your data based upon the real-world entities they represent or how the data or applications will be used. Do not merely mimic current implementations and existing database designs. This is especially important if the data is (or could be) used by many applications.

Create multiple interfaces to the data. Consider creating multiple ways to access data. For example, one might maintain a legacy data access approach while offering new access methods such as a web service that provides data mapped to an XML Schema as well as a web service that provides RDF and OWL tagged data. Anticipate that data access methods will likely change in the future and build in the flexibility to accommodate these changes and allow for "smarter" interfaces in the future. Modularize new pieces of information to make it easier to layer the solution.

Invest in Semantic Web Best Practices and Design Patterns. Participate in groups like the W3C Semantic Web Best Practices and Deployment Working Group.[20] Also, as you develop ontologies, lessons learned, and best practices, share them with the web community to help evolve the state of the practice.

Encourage DoD funded research in areas important to our DoD customers. This includes research in the areas of trust and security, a rules and logic layer for the Semantic Web, and approaches for enterprise ontology management.

9. Getting Started. For those looking for further background information on the Semantic Web, or who are ready to begin to leverage these technologies, here are some excellent resources.

An introduction to the Semantic Web was initially presented in an article that appeared in Scientific American in May 2001, titled "The Semantic Web – A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities" [BHL01]. This article provides a good overview and motivational content for a Semantic Web. Grigoris Antoniou and Frank van Harmelen, to name only two of the prominent researchers in the field, published the very useful book "A Semantic Web Primer" [AH04] in 2004. Finally, if you would like to understand how the Semantic Web relates to other technologies, you should read "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management" [DOS03], which nicely relates more mature and widely adopted technologies, such as XML and Web Services, with emerging technologies of the Semantic Web.

Ontology development tutorials are valuable resources to learn how to build and validate ontologies. A resource that we used is Stanford University's "Ontology Development 101: A Guide to Creating Your First Ontology" [NM01]. More formal discussion of the ontology engineering can be found in "Ontology Engineering" by Asuncion Gomez-Perez and Mariano Fernandez Lopez [GFC04]. This book does a nice

[20] http://www.w3.org/2001/sw/BestPractices/

job at discussing the ontology development process, the ontology life cycle, and methodologies for building ontologies.

Ontology development tools are emerging rapidly, so there are a number of tools to choose from. Those that are building ontologies should consider using Network Inference's Construct/Cerebra suite (ontology construction and management tool), Protégé (research tool by Stanford University), Jena (HP's open source custom Semantic Web application development framework), and Language Computing's Link Factory (ontology management system purposed for large scale ontologies).

Semantic Web standards are emerging. To ensure long term support and interoperability, it is important to be aware of and continue to follow progress of these standards. A focal web resource that provides, among others, standard specifications for Semantic Web technologies, is the World Wide Web Consortium web site.[21]

10. Conclusions. We believe Semantic Web technologies are here to stay and offer great potential to close the M2M semantic gap. Therefore, it is essential that we position ourselves and our customers to take advantage of these rapidly maturing technologies. We applaud the Government and MITRE investments to date and eagerly anticipate the advanced applications these technologies promise.

References

[AH04] Antoniou, Grigoris and Frank van Harmelen, "A Semantic Web Primer", The MIT Press, 2004.

[BHL01] Berners-Lee, T., J. Hendler, and O. Lassila, "The Semantic Web," Scientific American, vol. 284, no. 5, pp. 34–43, May 2001.

[DOS03] Daconta, Michael C., Leo J. Obrst and Kevin T. Smith, "The Semantic Web", Wiley Publishing, 2003.

[GFC04] Gómez-Peréz, Asunción, Mariano Frenández-López and Oscar Corcho, "Ontological Engineering", Springer-Verlag London, 2004.

[GG95] Guarino, N.; Giaretta, P. 1995. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: N. Mars, ed. Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing. IOS Press, Amsterdam: 25-32.

[NM01] Noy, N.F and D.L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology", Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

[OLW03] Obrst, Leo, Howard Liu, and Robert Way, "Ontologies for Corporate Web Applications", AI Magazine, Fall 2003.

[PSS04] Pulvermacher, Mary K., Suzette Stoutenburg and Salim Semy, "Netcentric Semantic Linking: An Approach for Enterprise Semantic Interoperability", Revision 1, MITRE Technical Report 04B0000075, The MITRE Corporation, October 2004, available online at: http://www.mitre.org/work/tech_papers/tech_papers_04/04_1174/.

[SLP03] Semy, Salim K, Mark H. Linderman, and Mary K. Pulvermacher, "Information Management Meets the Semantic Web", The MITRE Corporation, November 2003, available online at: http://www.mitre.org/work/tech_papers/tech_papers_03/03_1011/index.html.

[SP04] Sabbouh, Marwan and Mary K. Pulvermacher, "Wedding the Web: An Example of a Services and Semantics Marriage that Works," The MITRE Corporation, July 2004, available online at: http://www.mitre.org/work/tech_papers/tech_papers_04/sabbouh_web/.

[SPO04] Semy, Salim K., Mary K. Pulvermacher and Leo J. Obrst, "Toward the Use of an Upper Ontology for U.S. Government and U.S. Military Domains: An Evaluation", MITRE Technical Report 04B0000063, The MITRE Corporation, September 2004, available online at: http://www.mitre.org/work/tech_papers/tech_papers_04/04_0603/.

[SKS04] Semy, Salim K., Kevin N. Hetherington-Young, and Steven E. Frey, "Ontology Engineering: An Application Perspective", MITRE Technical Report, August 2004, available online at: http://www.mitre.org/work/tech_papers/tech_papers_04/04_0847/.

[21] http://www.w3.org/

Acknowledgements

We gratefully acknowledge the review and thoughtful comments on this paper by Dr. Arnie Rosenthal, Dr. Len Seligman, and Mr. C. J. Bohn. We would also like to thank Mr. Mickey Stinnet for his excellent formatting of this document.

About The Authors

Dr. Leo Obrst is a principal scientist at the MITRE Center for Innovative Computing and Informatics, where he leads the Information Semantics team. Formerly, he was Director of Ontological Engineering at VerticalNet.com, a department he formed to create ontologies in the product and service space to support Business-to-Business e-commerce. Leo's PhD is in theoretical linguistics with a concentration in formal semantics from the University of Texas-Austin. He has worked 20 years in computational linguistics, knowledge representation, and in the past nine years in ontological engineering. Leo was a member of the W3C Web Ontology Working Group that developed the Web Ontology Language OWL. He is a member of IEEE Standard Upper Ontology working group, the Semantic Web Services Initiative, and the Ontolog group, which he co-founded and which focuses on ontologies and business. He is a member of AAAI, ACL, LSA, and ACM. He is co-author (with Mike Daconta and Kevin Smith) of the recent book *The Semantic Web: The Future of XML, Web Services, and Knowledge Management.* John Wiley, Inc., June, 2003.

Mary Pulvermacher is a Senior Principal Engineer in the Center for Air Force Command and Control (C2) Systems at the MITRE Corporation in Colorado Springs, Colorado. With almost twenty years experience in Systems Engineering, Mary has most recently focused on leading Semantic Web research efforts sponsored by the MITRE Technology Program. Her work has included investigations into the use of ontologies, rule-based systems and intelligent software agents in the U.S. Military C2 Enterprise. Under her leadership, MITRE demonstrated that ontology linking across domains could be a powerful, extensible step toward semantic enterprise interoperability. Mary's work has also resulted in the identification of linkages between Semantic Web technologies and the USAF publish and subscribe information management approach, called Joint Battlespace Infosphere. Her teams have also assessed the value of using standard upper ontologies in DoD applications and identified key interoperability issues with the emerging rule language standard Rule Markup Language (RuleML). Mary is a Phi Beta Kappa graduate of the University of Wisconsin with a Bachelors of Science in Sociology and has a Masters of Computing Science from Texas A&M University.

Salim Semy is a Senior Software Systems Engineer in the Visualization, Collaboration & HCI Department of MITRE's Information Technology in the Air Force Center. Salim has been actively involved in multiple Semantic Web research initiatives at MITRE. The focus of this research is to investigate and apply Semantic Web technologies to address data interoperability challenges through semantic markup of data and machine to machine (M2M) interaction. In particular, Salim investigated the relationship between a publish and subscribe approach (such as the Joint Battlespace Infosphere (JBI)) and the Semantic Web, and applied semantic technology to develop a prototype knowledge base to facilitate collaboration across the Predictive Battlespace Awareness (PBA) community. Salim has also investigated the use of standard upper ontologies in the U.S. Military domain. Salim received an Honors B.Sc. in Computer Science from the University of Toronto in 2000 and is currently pursuing a Masters degree in Neurobiology at Boston University School of Medicine.

Suzette Stoutenburg is a Lead Information Systems Engineer at Air Force Command and Control (C2) Systems at the MITRE Corporation in Colorado Springs, Colorado. Suzette has over twenty years experience in Systems Engineering, with a formal background in Artificial Intelligence and rule-based systems. Most recently, Suzette has supported Mary Pulvermacher in her efforts to evaluate linking of ontologies across domains as a potential solution to semantic enterprise interoperability and integration. Also, Suzette examined the interoperability aspects of the emerging rule language standard for the Semantic Web, RuleML, and is now leading a research effort with Dr. Leo Obrst, entitled Toward a Standard Rule Language for Semantic Enterprise Integration. Suzette received a Bachelors Degree in Computer Science from Rowan State University in 1985 and a Masters Degree in Computer Science from National Technological University in 1989.