AFRL-IF-RS-TR-2007-49
Final Technical Report
March 2007

UNITED STATES AIR FORCE (USAF) SEMANTIC INTEROPERABILITYCAPABILITIES BASED ASSESSMENT AND TECHNOLOGY ROADMAP

Image Matters LLC

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

AFRL-IF-RS-TR-2007-49 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/

GENNADY R. STASKEVICH
Work Unit Manager

/s/

JAMES A. COLLINS, Deputy Chief
Advanced Computing Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
MAR 2007

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
Sep 06 – Jan 07

4. TITLE AND SUBTITLE
UNITED STATES AIR FORCE (USAF) SEMANTIC INTEROPERABILITY CAPABILITIES BASED ASSESSMENT AND TECHNOLOGY ROADMAP

5a. CONTRACT NUMBER
FA8750-06-C-0235

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
63789F

5d. PROJECT NUMBER
SEMI

5e. TASK NUMBER
06

5f. WORK UNIT NUMBER
02

6. AUTHOR(S)
Yaser Bishr
Timothy Eveleigh
Timothy Finin
James Hendler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Image Matters LLC
201 Loudoun St., SW
Leesburg VA 20175-3010

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/IFSE
525 Brooks Rd
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2007-49

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 07-067

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The primary purpose of this study is to assist the United States Air Force (USAF) to decide where and when to invest effort and resources to research, develop, and implement a family of technology believed to offer solutions to the problem of automated machine-to-machine (M2M) interoperability. The USAF has identified the automation of interoperability between machines as a key enabler of net-centric warfare. In this study we assert that semantic technologies are needed to enable true M2M interoperability yet at the same time acknowledge that semantic technology is at early stage of adoption in the industry and has not been widely adopted by the Department of Defense (DoD). Similarly, semantic interoperability (SI) has yet to enjoy rich support from doctrine, training, leadership, or as common material solutions.

15. SUBJECT TERMS
Net-centric, Semantics, Interoperability, Information Transformation, Information Management, Study

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UL	78	Gennady R. Staskevich
					19b. TELEPHONE NUMBER (*Include area code)*

Table of Contents

List of Tables

List of Figures

1 Executive Summary

The DoD's Net-Centric Data Strategy (DoD CIO, 2003) describes an "integrated approach for delivering a foundation for net-centricity." Recognizing that information is the fundamental enabler of net-centric warfare, the Department has consequently embarked on the wide-scale transformation of the way information is managed to "accelerate decision-making, improve joint warfighting, and create intelligence advantages." The Strategy's guidance details how programs and systems can transform themselves and their corporate information assets, from stovepiped point-to-point solutions into shared information resources that are visible, accessible, trustable, and understandable to all clients of the emerging Global Information Grid (GIG) – to include edge users.

Fully realized, the DoD's data strategy will enable the rapid, accurate, and autonomous machine-to-machine (M2M) exchange of critical decision support information. True semantic interoperability wherein machines exchange knowledge, and fully understand each other, will require the means to describe this knowledge in a common, consistent manner (i.e., machine-encoded facts and associated rules that represent the meaning and understanding of GIG resources, in domain context; represented as ontology), as well as require a host of supporting methods and services to support knowledge creation, discovery, quality of service, persistence, mediation, migration, and information assurance. The present governance model favored by OSD to achieve this realization asks communities of interest (COIs) to form and organize the compilation and definition of domain knowledge in the form of schema and ontologies. At this early stage in the implementation of the GIG, however, few COIs or programs have implemented the means to describe and exchange domain knowledge about GIG resources with the semantic richness necessary to fully enable the autonomous cross-domain machine-to-machine interoperability vision.

Over the last six years, the World Wide Web has also been undergoing a revolution in which its contents have become increasingly machine-processable instead of being only human-oriented. The Semantic Web is at the center of this revolution where an ever expanding group of technologies have been developed to enable computers to publish, discover, access, and understand networked resources (data and services) without human intervention. Several major technology firms now offer such technology in their mainstream software products (e.g., IBM, Yahoo, and Microsoft). By enabling semantic technologies in net-centric environments, these early adopters are beginning to reap the tremendous benefits to be gained from using semantic-based tools to arbitrate and mediate structures, meanings, and contexts within and between diverse communities of interest.

The USAF recognizes that semantic technology is becoming a reality and is a strong enabler of aerospace operations in the GIG. The USAF has developed this Semantic Interoperability Roadmap as a means to guide its adoption in the short and long terms, and as a decision aid for prudent investment in research that will hasten the adoption and full effectiveness of semantic interoperability technologies. The roadmap derives its strength from its use of the JCIDS Capability-Based Assessment approach to determine the requirements for semantic interoperability; describe functional capability gaps considering the current state of the practice of semantic technologies and net-centric enterprise policies and material solutions; and ultimately to explore DOTMLPF solutions to fill these gaps. The study shows that the effort and investment required to realize ubiquitous M2M interoperability is considerable, but that many of the policy and technology building blocks are now in place. The study also indicates where semantic technologies have matured and can now provide basic capabilities and where investment is still needed to realize the full potential of these technologies as enablers of information dominance.

2 Purpose of this Study

The primary purpose of this study is to assist the USAF decide where and when to invest effort and resources to research, develop, and implement a family of technology believed to offer solutions to the problem of automated machine-to-machine (M2M) interoperability. The USAF has identified the automation of interoperability between machines as a key enabler of net-centric warfare. In this study we assert that semantic technologies are needed to enable true M2M interoperability yet at the same time acknowledge that semantic technology is at early stage of adoption in the industry and has not been widely adopted by the DoD. Similarly semantic interoperability (SI) has yet to enjoy rich support from doctrine, training, leadership, or as common material solutions.

Because SI technology is relatively new and because it exhibits so much potential, a rational approach to its exploration and application must be developed so that it is not applied in too hasty a fashion, in ways where it is clearly inappropriate, or without due consideration of the lifecycle costs of its introduction. This rational consideration of what SI can do towards addressing USAF M2M interoperability problems is what this study and its Technology Investment Roadmap hope to perform.

Ancillary purposes of this study are to document the nature of requirements for M2M interoperability within the USAF, assess the maturity of present implementations of SI technology within the department, and identify functional gaps that must be addressed to enable future M2M interoperations.

3 Scope of this Study

The need to interoperate or communicate knowledge between functional elements of an enterprise exists in most if not all the functions, nodes, and stakeholders comprising the greater USAF enterprise. The systems-of-systems perspective further extends the USAF enterprise to the broader joint defense, defense department, and ultimately to the full government and the national and international commercial spheres. To survey the interoperability needs of all possible sender receiver pairs in this giant enterprise of enterprises is likely not a possible task. Instead, some measured scope must be defined to limit the investigation of cogent interoperability needs and instead focus on subsets that represent challenging problems where interoperability – or the lack of it – is often cited as a limiting factor to mission success. Predictably, these needs and the missions they represent also have the early stakeholder buy-in that we would hope could influence investment in the technologies we describe in this study.

Earlier studies, notably the USAF Science Advisory Board (AFSAB) study of Domain Integration (Appendix B, Ref M1) reached a similar conclusion and instead of focusing on all USAF-wide and department-wide domain knowledge exchanges, focused on a particular use case family where the dominant domain interaction:

1. Is central to a recognized core USAF mission
2. Has wide USAF stakeholder awareness
3. Enjoys department-wide awareness of the continued failure of existing technologies to deliver an effective and affordable solution
4. Has sustained and likely will sustain considerable investment to fix this shortfall

5. Has flight safety and lethal force considerations
6. Is a prime candidate for automation requiring M2M interaction

The use case family of concern involves aspects of the widely known and studied time-sensitive targeting operations. This use case family also involves sensors, shooters, decision makers, and knowledge and collection resources disposed in the well distributed and network-centric environment called the Global Information Grid (GIG).

The 2006 SI Working Group at the Minnowbrook Conference hosted by AFRL also chose this use case family, in part due to its consonance with the broader interoperability aims, setting, and scenarios of the ongoing Operational Information Management (OIM) program and its predecessor, the Joint Battlespace Infosphere (JBI) program.

The scope of this study inherits the influences of the AFSAB study and the OIM/JBI program in that it focuses on the M2M interactions necessary to perform core USAF air operations functionality – the strike use case. While we have scoped the study this way, the systems engineering methodology we use forces us to remain somewhat objective and in this context the functional needs of the TST engagement community to communicate between nodes and players is relatively similar (except maybe to accommodate a need for alacrity in a more harsh communications environment) to the needs of many other USAF communicators from personnel and supply systems, to finance and maintenance. For this reason, we believe that the Capability Based Assessment approach and the SI Technology Roadmap that was developed by this study are actually broadly extensible to the wider USAF functional community and to the greater DoD domain.

4 Semantic Enablement of M2M Interoperability

Semantic Interoperability is directly applicable to M2M interoperability in that it addresses one of the fundamental challenges of the GIG: ensuring that the vast federation of information producers and consumers (some determined, some opportunistic) can locate and understand the information they produce and consume. As this study will show, there are technological, political, and logistical challenges to achieving the ultimate goal of ubiquitous M2M interoperability within the GIG. This study will also show that M2M SI is possible within a relatively short time if investments and policies are coordinated and injected at the right time and in the right places.

Other technologies, notably messaging and relational databases have also tried to enable M2M interoperability with varying degrees of success. These efforts invariably attempt to establish and standardize a specific exchange format or data model. Where this approach has succeeded is where both parties to an exchange – sender and receiver – have full awareness of the meaning and structure of the terms that form the model; typically, however, this awareness must be encoded in software or hardware. This approach typifies the controlled and often brittle point-to-point interfaces that have historically dominated military communications. As the number and types of data and service providers in the Global Information Grid increases, the significance of data integration and interoperability attract greater attention to the need to open interfaces and share data. Semantic technologies also require the establishment of models to describe data but they are developed in a form where they can contain greater contextual nuance and are made machine readable.

As a precursor to this study, we make the following observations about the development and use of exchange models:

1. Independent and scattered development of data (and semantic) models will not lead to scalable interoperability solutions. Users of information systems spend substantial time

interpreting data and entering it into other applications. This is due to the fact that data developed by different people, for different purposes, with different constraints will have different structures and meanings. Therefore, independently developed data models, data dictionaries, and metadata each have unique perspectives, purposes, and constraints. These differences can lead to divergence.

2. Policies that encourage DoD-wide common data and semantic model development (e.g., universal data models and "upper" ontologies) do not scale and have led to large, unmanageable modeling efforts. Broader coordination across DoD can lead to interoperability within larger domains, but cannot scale to the levels of a large multi-faceted enterprise like the objective GIG.

3. Interoperability efforts have largely been focused on syntactic rather than semantic aspects.

It is useful to note that these assertions apply to both semantic and to non-semantic interoperability technologies (e.g., relational databases and message traffic). For example, no common master database schema or master message format has ever evolved despite numerous attempts to define such. Similarly a profusion of localized databases and message formats has only further complicated interoperability. Database schemas and message formats provide a syntax that allows information to be labeled and formatted but in most instances the receiver of information arriving from databases or delimited message traffic must already possess or gain additional information about the context of the arriving data to be able to fully understand its content. Historically, this has been a challenge for machines. Semantic technology differs in that the context for and relationships between information elements can also be relayed and this understanding is available in a machine-processable form. This distinction underlies the difference between syntactically described information and semantically described knowledge. By *semantics* we refer to the meaning of phenomenology as it is represented in computer machines whereas *syntax* refers only to the form or representation of the information.

4.1 Semantics and Semantic Interoperability Defined

SI between systems can be defined in terms of information that flows between them. Semantics is defined as the meaning or relationship of meanings of terms and expressions within a system. SI can therefore be defined as the ability of information to flow between systems on the basis of shared, pre-established, and negotiated meanings of terms and expressions such that information is accurately and automatically interpreted by the receiving system.

Shared understanding of meanings between systems is a necessary condition for information to flow between systems. This shared understanding is only possible when there are regularities as well as constraints on these regularities within and across these systems. The term regularity refers to an observed pattern in the world. Wrightson 2001 acknowledges that each individual has a different view of the world, called a Scheme If Individuation, or SOI. Within these SOI, regularities occur - and these regularities are a formalization of a common knowledge. The ultimate goal in M2M, therefore, is to represent these SOIs and for information to seamlessly flow between them.

SI is not, and should not, be limited only to data interoperability. To fully enable M2M interoperability SI must also encompass all aspects of machine-to-machine communications, including services, security, and quality of services.

4.2 Semantic Interoperability Issues

While semantic technology offers the promise of improving interoperability by formally exchanging meaning between communicating actors, it also introduces several issues that must be addressed to make it effective in the USAF GIG M2M use case.

4.2.1 Semantic Heterogeneity

The most challenging issue that faces SI is semantic heterogeneity. Since semantics deals with human interpretations according to understandings of the world, it is therefore context dependent. Different interpretations of data lead to semantic heterogeneity. A database is considered consistent if all its content satisfies all user defined consistency constraints. Consistency constraints are usually derived from semantics of data items in the application domain. For example, highways should not intersect with rivers unless there is a bridge. A lack of semantic consistency in turn can limit interoperability between systems if it is not accommodated. Thus, technology that can exploit semantics *and* context is crucial to achieving semantic interoperability.

In the context of M2M interoperability, we define SI as the ability of communicating agents to understand each other with a guaranteed accuracy. This is a requirement for complete semantic integration in which the intended models of both agents are mutually understood and consistent, that is, all the inferences that hold for one agent, should also hold when translated into the other agent's ontology. In the context of SI, ontology is defined as machine readable specifications of conceptualization of real world phenomenology.

An ontology is a tuple: $O = (C, R, \leq, \perp, |, \sigma, I)$ where:

> *C is finite set of concept symbols;*
>
> *R is a finite set of relation symbols;*
>
> *$\leq$ is a reflexive, transitive and anti-symmetric relation on C (a partial order);*
>
> *$\perp$ is a symmetric and irreflexisve relation on C (disjointness);*
>
> *| is a symmetric relation on C (coverage)*
>
> *$\sigma : R \Delta C^{/}$ is the function assigning to each relation symbol its arity;*
>
> *the functor $(-)^{/}$ sends a set C to the set of finite tuples whose elements are in C; and*
>
> *I are instances that belong to C.*

From this definition, we can conclude that heterogeneity between machines can arise from any or all of the elements of the ontology tuple, i.e., differences in labels of concepts, relationships between symbols, classification, and/or constraints defined on C with respect to R.

M2M interoperability is defined as the ability of systems to independently and yet transparently communicate at all levels of the technology stack, specifically the information content layer. Heterogeneity mainly arises as a result of differences between any layer of the technology stack including but not limited to the: operating system, network protocol, application interfaces, and information content. In this section we will only focus our discussion on heterogeneity as a result of differences between information content in different systems. Semantic heterogeneity is widely regarded as the most significant obstacle to the successful sharing and exchange of data as it can limit the reuse and sharing of data between agencies or across different applications within the same agency. Data development efforts are often limited in scope, which severely limits the ability to apply available data to effective, multi-discipline decision-making. Worse still, users must often perform laborious data conversion tasks, translating data sources developed for one domain into a new domain, without understanding the limitations of using the data in the target domain. Of equal concern is the tremendous difficulty in integrating data obtained by different organizations.

Semantic heterogeneity can take many forms, including but not limited to, differences in naming, scaling, confounding (e.g., "real-time news" does it mean five minutes delay, ten minutes delay or no delay). The classes of conflict that arise from semantic heterogeneity are well documented in literature (Kim 1991). Semantic heterogeneity can also exist in geometric descriptions of features as a result of merging (integrating) different data sources. For example, suppose that within one domain a line is defined by two points, and in another domain, a point is defined by the intersection of two lines. Merging these models results in a circular reference to the resultant tree model and is therefore infinite. Users from different communities, who share their data, are likely to share interest in a common understanding of the real world.

The wide-spread use of ontologies by diverse communities and in a variety of applications is a commonality in today's knowledge sharing efforts. They are the backbone for semantically rich information sharing, a prerequisite for knowledge sharing. As systems become more distributed and disparate, within and across organizational boundaries, there is not only a need to preserve the meaning of concepts used in everyday transactions of information sharing but also the mappings between ontologies that describe those concepts.

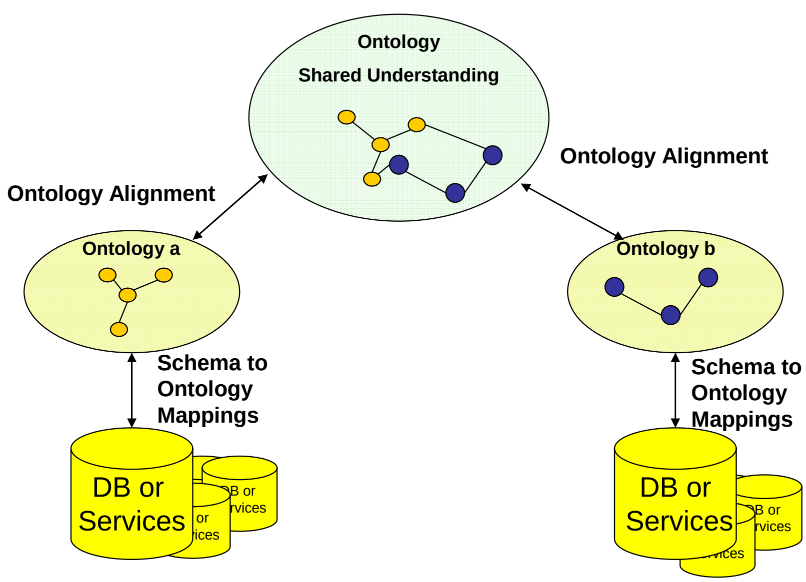

Figure 1, SI within and across communities

Figure 1 depicts the technical challenges involved to accommodate semantic heterogeneity. Specifically, to develop a shared understanding between two communities of interest or domain systems, the local information content of each domain must be mapped to a domain ontology and an alignment between the two ontologies established. These two key concepts are now described:

- **Schema to Ontology Mapping** (S-O Mapping): This closely matches SI within communities of interest (COIs) by mapping the local ontology of the domain to the underlying databases and services. The result is an ontology that is populated by instances retrieved from the underlying databases and services. We call the resulting populated ontology a Knowledgebase.

- **Ontology Alignment**: This closely matches SI across or between COIs and involves mapping the domain ontologies to an ontology that represents a shared understanding of the underlying systems.

4.2.2 Semantic Richness

A second important consideration with SI is semantic-richness[1] or the degree to which information is provided semantic description. It is generally acknowledged that semantic richness is limited or lacking in most IT environments. Semantic properties are generally hardwired into applications, thus limiting flexibility and interoperability, or they are missing altogether, thereby encumbering users with tedious, manual processing tasks. The lack of sufficient semantic information can lead to the following problems:

1. Semantic heterogeneity (i.e., differences in meaning and significance) as described in the previous section occurs inhibiting data sharing
2. Users must perform tedious tasks to employ semantic-limited data (where the tedious nature of the tasks stem from the lack of semantics)
3. Users employ semantic-limited data inappropriately (in an application for which they are not intended)
4. Users misinterpret the meaning of semantic-limited data while using it in an application

These problems have serious and costly consequences for any enterprise. For these reasons, any application of SI technology will have to consider the extent to which shareable resources are semantically annotated so as to minimize the need for additional transformation, misunderstanding, or human intervention. This richness factor may not necessarily be addressable from a pure technology perspective as ultimately humans and programs must decide (or be told) to add semantic content to their information assets. Thankfully, this is the focus of the DoD’s Net-centric Data Strategy. As we will see, however, implementation of this policy has been piecemeal and has not yet resulted in the semantic depth necessary within most application domains to assure full M2M interoperability.

Taken together, semantic heterogeneity and limited semantic richness impact the following elements of the enterprise:

1. **Flexible querying of databases, schema integration, and automatic data translation** – the lack of semantics at the database level greatly limits the interoperation and integration of distributed databases under common application frameworks within an information community
2. **Service automation** – limited semantics and semantic heterogeneity inhibit service automation by requiring users to supply missing information, overcome interoperable connections, etc.

[1] By “semantic-richness” we mean the concepts that capture and convey the full meaning and significance of phenomena and the understanding of how these phenomena behave. For an information system, semantics would describe how (what, where, why, when, and how) actors, their roles, data, and business processes are involved in an information processing environment that transforms raw data into higher forms of user information, understanding, and knowledge.

3. **Service personalization** – the lack of semantics at the user level limits service customization and configuration, and in particular, the ability to tailor services for a given user and user situation
4. **Application extensibility** – the lack of semantics at the business process level greatly limits the interoperation and integration of distributed applications and business processes.

4.2.3 Implementation Issues

Applying semantic technologies to enable M2M interoperability also brings logistic issues that must be addressed. Considering the full information lifecycle, these include but are not limited to: providing means to generate, maintain and control semantic ontologies; grounding them to information assets and services; making them discoverable; making them persist; assuring their quality; and to making them operationally available when needed. Similarly, SI technologies will have to support numerous department policies and standards such as information assurance and the use of joint terminology. Semantic technologies may also have to accommodate non-semantically-enabled legacy systems. In short, semantic technologies introduce a host of new requirements that must be understood and accommodated before they can be used in an operational USAF setting.

This study attempts to systematically investigate not only the basic semantic capabilities needed to implement domain interoperation, but also to explore the broader class of other derived and implied functions needed to operationalize the technology. For this, we applied the systematic Joint Capabilities Integration and Development System's (JCIDS) Capabilities Based Assessment methodology.

5 JCIDS Capabilities Based Assessment Approach to Roadmap Generation

CJSCI 3170.01D (JCS, 2004) describes the Joint Capabilities Integration and Development System (JCIDS) methodology that implements the current acquisition doctrine of the DoD. JCIDS suggests a series of analytical steps called a Capabilities Based Assessment (CBA) that has the intent of grounding proposed acquisition actions and problem solving capabilities to acknowledged functional needs of the DoD while considering the capabilities of available technology and existing programs. JCIDS also establishes a review process whereby prospective capability analysis and solutions are subjected to critical functional review prior to capability production. JCIDS encourages good systems engineering practices such as the full investigation of functional capability needs before suggesting candidate solutions. Similarly, JCIDS encourages the consideration of multiple solution alternatives which must by definition consider doctrinal, organizational, training, leadership, personnel, and facility means to achieve capabilities in addition to material solutions. In this sense, JCIDS develops DOTMLPF recommendations for how to solve functional problems within the DoD.

We chose to use the JCIDS model to structure this study so that the study would not become too focused on one area of technology, miss entirely the opportunity to capture traceability between technology and required capabilities, or fail to consider non-material solutions. We also felt that a JCIDS approach would take a positive step towards making recommended solutions “acquisition-ready” in that they would have at least in part already been exposed to JCIDS rigor. Since this study amounts to a Capabilities-Based Assessment of SI needs and technology, the study is conceptually interoperable with and directly supports the Joint Requirements Oversight Council’s Functional Capability Boards.

Finally, by bringing JCIDS rigor early in AFRL’s proposed outyear SI research program, we hope to establish a common, repeatable, and acquisition-ready process to support the remainder of the program. We feel that this will be an important step in ensuring that AFRL’s overall SI investment delivers an internally consistent and well-balanced product featuring rigorous traceability back to originating network-centric capability requirements. Any capability acquisition that results from the outyear research activities should thus be provided with a shortened path to “Milestone A” consideration, as much of the preparatory analysis will have been accomplished.

5.1 Capabilities Based Assessment

To be useful and credible, this study and its roadmap development activities must perform several key activities. These include:

- Investigate current and future USAF requirements for net-centric interoperability
- Investigate the current state of the science and state of the art in SI technologies
- Determine how semantic technologies might best be applied to solving military interoperability problems
- Identify technology shortfalls, bottlenecks and barriers, and determine the most prudent future research investments based on research needs and priorities
- Recommend future research and development topics, objectives, needs, priorities, risks, and approaches for achieving SI within and between USAF systems

These activities are directly parallel to the Functional Capabilities Analysis processes within JCIDS that "identify prioritized capability gaps and integrated joint DOTMLPF and policy approaches (materiel and non-materiel) to resolve those gaps." Re-written in JCIDS terminology, the activities become:

- Perform a Functional Area Analysis (FAA) that will collect, derive, and structure current and future capability requirements for USAF net-centric SI
- Conduct a Functional Needs Analysis (FNA) that assesses the current state of the science and state of the practice in semantic technologies and programs, as they may contribute to achieving SI for USAF systems
    - FNA determines how semantic technologies might best be applied to solving military interoperability problems
    - FNA identifies technology shortfalls, bottlenecks and barriers, and determine the most prudent future research investments based on research needs and priorities
- Conduct a Functional Solution Analysis (FSA) to recommend a future research and development agenda and plans for achieving SI within and between USAF systems.

Figure 2 shows these basic steps.

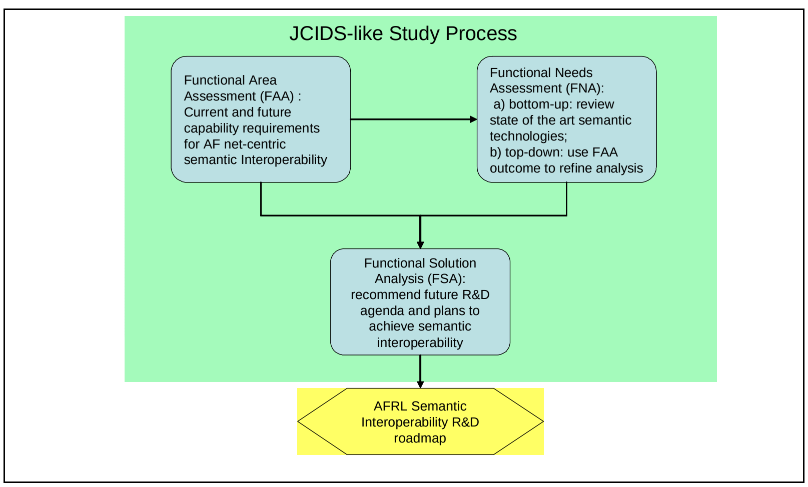

Figure 2, Application of JCIDS to develop SI Roadmap

5.1.1 FAA Purpose

This stage of the study investigates capability requirements for M2M SI and develops a functional architecture model for SI that supports the USAF role in the GIG. This model will propel and structure both the subsequent FNA and FSA stages.

5.1.2 FNA Purpose

The FNA stage of the study continues the systems engineering process started by the FAA stage and assesses the ability of current and emergent SI technologies and programs to deliver the capabilities the FAA stage identified under the full range of expected network-centric operating conditions and to designated measures of effectiveness.

5.1.3 FSA Purpose

The FSA examines and/or recommends potential DOTMLPF and policy approaches to solving (or mitigating) the capability SI gaps identified in the FNA. The FSA stage further recommends a future research and development agenda for achieving SI within and between USAF systems, with M2M interoperability as the ultimate goal. This agenda is in the form of a roadmap that identifies future directions and topics for research and development, as well as goals, objectives, needs, priorities, and risks. The FSA stage also defines the recommended methods and procedures to achieving the plan in terms of short-term studies, long-term studies, prototyping activities, operational tests, and so on. The completed FSA documents the capability gaps and recommended research approaches and includes traceability back to the originating capability requirements and issues. Finally, the FSA also highlights ways other than material solutions that the USAF might address SI capability gaps (such as leveraging the COI governance model or promulgating key standards).

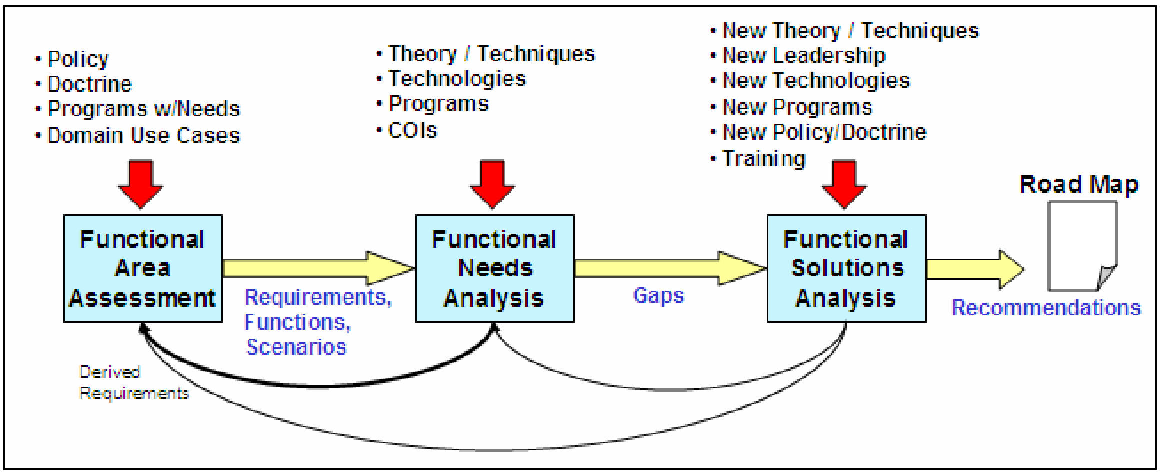

Figure 3, Interaction between the analysis stages.

5.2 Interaction and Flow Between the Analyses

To some degree we subvert the full JCIDS systems engineering process in this study in that we are acting under the overall presumption that SI technologies represent a good part of the solution set for USAF M2M interoperability challenges. With this perspective, the CBA must also investigate requirements that are implied by this solution choice (see Section 4.2.3). For example, by concluding that semantic ontologies are needed to capture and represent structured knowledge and rules for a particular domain, we must be willing to recognize that adoption of such a technology begets the need for supporting capabilities – in this example, a means to define, make persistent, and use the aforementioned ontologies. In a similar vein, we know that the DoD and the USAF have already written policy and developed social and material means to implement SI technology (e.g., the COI concept and the GIG Enterprise Services (GIG-ES) program). Recognizing via the FNA that these are proposed solutions we again see that these solutions also beget requirements. Taken as a whole, this application of CBA to SI technology must therefore assess requirements from three sources: the stated requirements and desired capabilities of the USAF and the doctrine and policy mandates of the executive branch and the USAF; the requirements implied by existing policy and solution implementations; and the requirements implied by suggesting the adoption of specific technologies to fill remaining gaps. This backward flow of requirements is shown in Figure 3.

6 Functional Area Analysis

6.1 Method

Our approach to the FAA was to survey recent DoD and USAF policy concerning interoperability, information exchange, and net-centric operations (The "P" series documents in Appendix B lists the various sources we considered in this analysis). Reading through these documents we collected what amounted to performance, technical, and functional requirements noted (or in some cases implied) by the documents. As these requirements were collected, we gradually emerged a detailed structure of requirements which we subsequently clustered into taxonomy (Appendix C). Finally, we collapsed the detailed functional requirements taxonomy into a simplified functional architecture which we used to simplify and structure the FNA and FSA. We preserved linkages between successive categories and back to the original source documents. As the analysis progressed into the FNA stages, we examined the implications of the technical and non-technical solutions and programs presently promulgated by the department (these are the "C," "D," "M," and "U" series sources in Appendix B.) The implications of these sources were also placed in the emerging taxonomy. Similarly, as we conducted the FSA stage, we again assessed the implications of DOTMLPF solutions recommended to close identified interoperability capability gaps. Taken together, the FAA stage generated a comprehensive view of the requirements for SI and the lifecycle implications of adopting it as a solution choice.

6.2 Findings

The Defense Acquisition University (DAU) Guide (Reference P21) states the *GIG Overarching Policy Directive*'s (P11) vision to "empower users through easy access to information anytime and anyplace, under any conditions, with attendant security." The Guide further notes that this vision will require a comprehensive information capability that is "global, robust, survivable, maintainable, interoperable, secure, reliable, and user-driven." The Guide also expounds the GIG goals for enterprise and community data, which includes visibility, accessibility, understandability, trustability, interoperability, and the ability to be responsive to user needs. This level of guidance was useful in structuring the functional requirements taxonomy (see Appendix C) of required semantic capabilities and to decompose implications for, and expectations of, SI technology.

The DoD Chief Information Officer's (CIO) guidance on transforming the present DoD IT enterprise to implement the GIG via the *Global Information Grid Architecture* (P25), the *Net-Centric Operations and Warfare (NCOW) Reference Model* (P41), and OSD's memorandum *GIG Enterprise Services: Core Enterprise Services Implementation* (M7) were also useful sources from which we collected capability requirements for SI. These sources, for example, identify major functional activity blocks including the requirements associated with interacting with net-centric services, user/entity services, providing net-centric service capabilities (core, COI, and enterprise control), provision of resource service requests, and enterprise information environment management components.

Regarding transforming the DoD enterprise to a network-centric environment, the *Net-Centric Enterprise Services Strategy* (P20), the *DoD Net-Centric Data Strategy* (P19), and the *DoD Information Assurance (IA) Strategy* (P15) each describe capabilities required to achieve four key GIG attributes: reach, richness, agility, and assurance. The DoD IT Standards Registry (DISR) contains a wealth of primarily commercial standards which further expound on and characterized these attributes. The attributes are reiterated by DoD Directive 4630.5 (P7), *Interoperability and Supportability of Information Technology (IT) and National Security*

*Systems (NSS)***.** This document was a very useful source to find and characterize SI capability and performance requirements, as it introduces the *Net-Ready Key Performance Parameter (NR-KPP).* NR-KPP provides performance requirements in the form of measurable, testable, or calculable characteristics and performance metrics required for the timely, accurate, and complete exchange and use of information. The NR-KPP also reaffirms the NCOW and further exposes applicable GIG *Key Interface Profiles (KIP)*, DOD information assurance requirements, and supporting integrated architecture products required to assess information exchange and use for given capabilities.

We found DoD's *Net-Centric Data Strategy* (P20), mentioned above, a very useful source as it describes the requirements for inputting and sharing data and metadata, and forming dynamic communities to develop semantic content (the COIs). This document was also useful in that it identified specific GIG architecture attributes such as data centricity, only handling information once, smart pull, posting in parallel, and application diversity. These are also provided in OSD NII CIO's *Net-Centric Attributes List* (P39) and the popular *Net-Centric Checklist* (P40). The *Net-Centric Information Assurance Strategy*, also noted above, was useful in that it describes the DoD's strategy for integration of information assurance into the global, net-centric information environment and thus opportunities for the insertion of SI technologies. Additional detailed information assurance certification and accreditation were derived from the 8500 Series documents and DoD Instruction 5200.40. Useful detail on data asset visibility was derived from OSD NII's memorandum: *Net-Centric Data Strategy: Visibility - Tagging and Advertising Data Assets with Discovery Metadata* (P42), *Department of Defense Discovery Metadata Specification (*D6), and DoD Directive 8320.2, *Data Sharing in a Net-Centric Department of Defense* (P14).

DISA's GIG ES Capability Development Document (M12) and its supporting Network-Centric Enterprise Services (NCES) program documents (M13-M25) were a great source for deriving semantic technology functional categories as they focused on the nine core enterprise services provided by the NCES Program. These categories include: application, collaboration, discovery, enterprise service management, information assurance/security, mediation, messaging, storage, and user assistance. This document set is tightly coupled with the NCOW RM.

Several sources, notably JROCM 199-04 (P43), CJCS Memorandum CM-2040-04 *Assignment of Warfighting Mission Area Responsibilities to Support GIG ES* (P44), and DoD Directive 8320 (P14) and its 2006 guidance annex (P13) expressed the Department's desire to implement network centricity and the establishment of domain ontologies and services via the COI process. The COI governance model process (as detailed in four COI guidance documents (C4-C7)), in addition to encouraging the wide scale compilation of domain semantic information and representations has introduced some additional challenges for SI technologies, notably the need to be able to interoperate between a profusion of vertical domain ontologies, each with controlled vocabularies and partial adoption of GIG enterprise standards (we will discuss this issue with more detail in the FNA).

The most recent intelligence community guidance comes from the newly released (Nov 2006) *Information Sharing Environment (ISE) Implementation Plan* (P32, P34) in which the Office of the Director of National Intelligence (ODNI) reaffirms DoD's data sharing plans and the mandates of PL 108-508 *Intelligence Reform and Terrorism Prevention Act of 2004* (P33) and EO 13338 *Further Strengthening the Sharing of Terrorism Information to Protect Americans* (P24). ODNI plans in early 2007 to release further guidance on implementing the ISE across government departments. The authors do not expect this prevailing guidance to in any way dissuade semantic solutions to cross program interoperability.

Appendix C presents a summary of the evolved SI functional taxonomy. The first six major categories were derived from the six major tenets of the DoD Net-Centric Data Strategy expressed in DoD 8320.02 *Data Sharing in a Net-Centric Department of Defense*. We have grouped high-level requirement categories within each of these divisions as well as specify a set pertaining to management tools and other policy-based requirements not covered by the other categories. This fundamentally is the product of the FAA.

6.3 Functional Architecture

For the purpose of segmenting the roadmap into functional categories and to simplify the requirement structure developed so far, we constructed a simple functional architecture that we use in the FNA/FSA to relate aggregated functional requirements to the existing and potential solutions that populate the roadmap. Figure 4 depicts the basic SI functional architecture. The arrows in the architecture indicate dependencies between its elements. For example, Migration depends on Persistence. Realizing these dependencies will assist the government in prioritizing its R&D investment strategy.

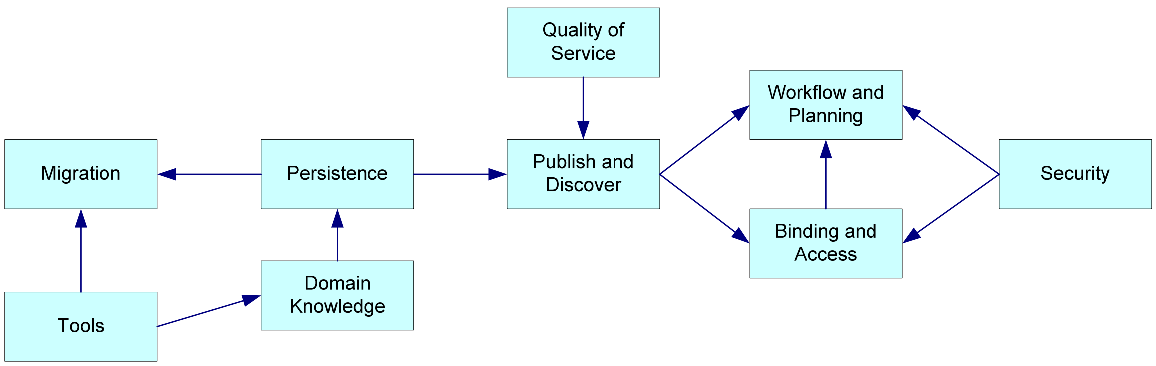

Figure 4, SI Architecture functional elements

We will now describe the nature of each of these elements.

6.3.1 Domain Knowledge

The Domain Knowledge functional requirement category arises from the understanding that semantic technologies encode and represent domain knowledge in a form that machines can interpret. Domain knowledge functional requirements are those that state the need for means to encapsulate domain knowledge in a semantic representation via some combination of manual, assisted, or automated means. Domain Knowledge functional requirements also to some degree express the need to address situations where the domain is assumed and thus knowledge is exchanged with tacit assumptions about the domain context. Similarly, interactions where no such assumptions are made and communication must occur in an explicit fashion must also be supported.

A number of requirements in the Domain Knowledge category concerned making provisions for representing the understanding of specific concepts. These included space and time, resource identity within a domain, classification level and releasability of information, and user intent. Other requirements concerned the desire to handle uncertainty and ambiguity or to assist with the identification of semantic contradictions. Understandability factors also drove the requirements to convey the role of exchanged information and of its receivers and senders. The FAA and FNA further identified the requirement for means to handle or otherwise accommodate controlled and uncontrolled vocabularies.

A common requirement noted was the requirement to tag content. While the concept of semantic tagging appears well known, little normative guidance is suggested for what should be tagged, how items should be tagged, strategies that suggest a link between tagging and context or workflow, or strategies or methods for detailing object classification or managing complexity.

Further requirements, both stated and implied express the need to be able to handle internationalization of units and terminology and to uniquely identify, serialize, or name objects within specific domains.

6.3.2 Publish and Discover

This category of functional requirements describes the need to make semantic content and services available to planned and potential consumers (the publish element) or to those seeking content and services (the subscribe element). Standing requirements were found to exist for the publishing and discovering of schemas and services but little discussion was found concerning the need to publish and discover higher-order semantic content such as ontologies and the mappings and correspondences between ontologies and schemas. These we inferred. The majority of exposed publish and discover requirements seem to concern making registries and catalogues accommodate particular elements such as DDMS/NCOW protocol and to handle issues like IPv4 vs IPv6. Similarly, requirements exist that stipulate that registries and catalogues indicate attribute knowledge like public vs. private intent, support for particular domains and edge devices, support enterprise searches, and to explicitly identify their contents and capabilities (e.g., publishing service descriptions and data dictionaries).

6.3.3 Tools

By the Tools functional category we mean the aggregate requirement for tools to support and enable the semantic enterprise. Key semantic tool categories include capabilities that:

- Establish ontological representations from existing domain knowledge
- Manage and version control ontologies and schemas
- Utilize knowledge engineering patters and best practices
- Visualize semantic contents
- Allow the combination and/or differencing or comparison of ontologies
- Assist in the conversion of schemas and vocabularies into machine readable ontologies
- Simplify the definition of mappings between ontologies and between ontologies and schema

The FAA and FNA also located both implied and explicit requirements for tools that support the use of SI technologies in the broader enterprise. Specifically sought were tools to generate, self-identify, manage, and where needed, dissolve semantic elements and instances within the infosphere.

As Figure 4 shows, the Tools category most directly influences the Migration and Domain Knowledge categories as both require the definition and maintenance of knowledge structures and linkages.

Almost all requirements for SI Tools were implied by the choice of semantic technology as a solution in the FNA and FSA stages; the need for semantic tools was rarely if ever mentioned in policy but is affirmed by several present programmatic efforts and frequently identified in lessons learned activities as a shortfall.

6.3.4 Persistence

Requirements for this category of functional requirements are expressed in a number of ways. Persistence requirements deal with the need to store or, in some way, persist semantic structures. These requirements describe the need to store both the knowledge described semantically but also the knowledge of the semantic representations of knowledge. Persistence is strongly related to the Migration requirement category in that the recall of semantic information from repositories is often accompanied with the need to migrate it to other representations and that persistence often plays a role in the migration process. Persistence is related to the Domain Knowledge category in that once domain knowledge is captured, it must be somehow persisted to be useful. Persistence also naturally leads to the Publish and Discover requirements category as once persistence is available, it must support means to add or withdraw contents.

6.3.5 Migration

The Migration functional category represents those functions that are needed to transport, convert, or otherwise exchange semantic elements from one representation to another. Several types of migration are noted or implied in the surveyed documents, these include:

- Migration between representations of similar expressiveness
- Migration between instances or versions of the same representation
- Migration between structures representing different levels of abstraction or expressiveness

The first of these types details cases where interoperability is desired between representations using the same basic semantic depth (expressiveness) or syntax. Examples here include schema-to-schema or ontology-to-ontology migration where the expressiveness of the source and destination are at rough parity. As the FNA will show, current state of the practice within DoD has focused on schema-to-schema migration (typically via stylesheets and simple mediation).

The second type of desired migration details those cases where the interaction concerns variations or instances of the same knowledge structure. A classic case of this occurs when an older and a newer version of a schema or ontology must interoperate. The challenge with this form of migration is in knowing precisely the implications of change in a schema or ontology and managing migration with this knowledge. Implied requirements were noted for the ability to version mark and/or uniquely identify schemas and ontologies to permit this form of migration but the deeper implied requirement for this case (and likely the other two cases) is for some form or reasoned mediation.

The third type of migration deals with requirements to migrate information between structures of varying expressiveness. An example of this is the need to migrate relatively shallow representations (e.g., an XML schema) and a richer abstraction (e.g., an RDF-described ontology with business rules).

These three cases imply requirements for mediation between representations likely involving some forms of reasoning and the need to develop and share concept mapping between representations. It is likely that context will play a role in these interactions and thus may also have to be described and or considered in the interoperation. Similarly implied is the need to compare representations and in some cases fuse or infer concepts to bridge representations.

There is little mention of the Migration functional category within policy or doctrine, rather it is implied by the choice of semantics as an solution to the M2M interoperability problem. As such, migration presents a formidable, yet poorly exposed functional need. As more domains develop semantic artifacts and department pressure to use semantic artifacts to enable cross domain interoperability increases, the challenges of bridging dissimilar representations, levels of expressiveness, and versions of representations will become paramount and technical, policy, organizational, and other solutions will be sought.

6.3.6 Security

Many of the policy documents detail requirements for security, releasability, and information assurance, essentially to support a defense in depth approach. In particular the FAA revealed requirements to implement specific security technologies such as PKI, XML encryption, and XML signature. Similarly, we noted requirements for security assertion, and the marking of content for releasability and handling caveats. More on the functional side, semantic technology was suggested as a means to support the adjudication of access rights, user authentication, and the operation of security guards. Finally, requirements were expressed for an ability to extend a security context and to mediate or broker access to marked content.

It is of little surprise that security requirements figure prominently considering much of the expected interoperation concerns secure content. Furthermore, the promise of easy discoverability and support to unanticipated users in many senses goes against decades of security practice.

6.3.7 Quality of Service

Quality of Service (QoS) requirements deal with a number of factors concerning the availability and reliability of semantic technology services. These include the support for edge users, interoperation in low bandwidth environments, and support for discontinuous operations (allowing intermittent connection with the enterprise). QoS requirements also stipulate the need to honor Service Level Agreements and the use of pedigrees, provenance, and other means to establish or convey trust between parties to a semantic interchange. Further QoS requirements detail the need to be able to validate schema and ontologies, identify possible contradictions between terms, and to ameliorate the impact of changes to semantic elements. A final category of requirements in this group deals more specifically with the exact exchange of semantic content and lists checksums, acknowledgement, and content verification as requirements. While in Figure 4 we show the QoS requirement category primarily influencing the Publish and Discovery functional category, it, like the Security category, really applies to all the other categories.

6.3.8 Workflow and Planning

The category Workflow and Planning category details requirements for the scheduling of semantic service invocations, the chaining of services together to perform tasks or support transformations, and the means to prioritize information delivery. These requirements were largely implied during the FNA/FSA vs. stipulated in doctrine and policy.

6.3.9 Binding and Access

This category covers requirements for binding and access to semantic resources. Much like the Publish and Discover requirements (and indeed supporting them) binding and access requirements detail requirements to actually locate and deliver the resources and services referenced in catalogs. In some cases, this may actually require the composition or assembly of content from multiple sources or the decision to migrate knowledge between forms to satisfy the needs of a requestor. Also implied in the Binding and Access category are requirements to select from multiple possible sources considering the context of the requestor; this could have releasability, quality of service, accuracy, permission, or domain specific considerations.

Binding and access to services also takes on this complexity. Again, these requirements were largely implied by the technology choice vs. expressed in doctrine or policy.

6.3.10 Other

The remaining requirements tended to be policy and governance requirements that stipulated compliance to various regulations, registries, reference models, and checklists. We did not consider these requirements in the FNA and FSA stages as most these requirements can be met with either documentation or staff action.

7 Functional Needs Analysis

7.1 Method

The centerpiece of this stage is the assessment of the current state of the science and state of the practice in semantic technologies as they contribute to achieving SI for USAF systems. The purpose of the stage is to identify and categorize current and emergent technologies (e.g., Semantic Web ontologies, security and privacy policies, logical inference engines, trust and social networks) and certain recent exemplar programs (e.g., Cursor on Target, Common Battlespace Object, FIOP Network Based Services, Cross-Service Weapon Target Pairing, Common Mission Definition), and to develop a set of matrices that compare existing state of the art technologies against the list of capability requirements developed during the FAA stage.

In addition to assessing semantic technologies, the FNA stage identifies semantic technology shortfalls/gaps, bottlenecks, and barriers that require solutions and indicate the timeframe in which solutions are needed. This knowledge is crucial to the FSA stage; were we recommend the SI research roadmap. The FNA stage also identifies redundancies in capabilities that may reflect inefficiencies and if needed, contrasts technologies that appear duplicative. The FNA stage prioritizes the gaps it identifies and further defines and refines the capabilities requirements architecture.

While semantic technologies are the focus of this study, we also believe that SI will be enabled by more than just technologies. As the OMG and COI processes have shown, achieving network-centric interoperability using semantic technology also requires consideration of non-material approaches such as polices, training, leadership, management, and new development methodologies. For this reason, the FNA assesses the entire range of DOTMLPF and policy, as an inherent part of defining capability needs.

7.2 Findings

Our FNA findings are broken into two sections, the first focuses on the current state of semantic technology research and development, the second on DoD progress towards implementing network-centric data sharing policy.

7.2.1 Current State of Semantic Technology R&D

We investigated a broad representation of SI research programs and technologies; Appendix D provides a list of these. Using the Technology Readiness Levels shown in Table 1, we estimated the readiness of these technologies with respect to their usability in operational environments.

Table 1 Technology Readiness Levels

Technology Readiness Level	Description
1. Basic principles observed and reported	Scientific research begins to be translated into applied research and development.
2. Technology concept	Invention begins. Once basic principles are observed, practical applications can be invented.
3. Analytical and experimental critical function	Active research and development is initiated.
4. Component validation in a laboratory environment	Basic technological components are integrated to establish that the pieces will work together. This is relatively “low fidelity” compared to the eventual system.
5. Component validation in a relevant environnent	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements.

6. System/subsystem prototype demonstration in a relevant environment	Representative model or prototype system, which is well beyond the breadboard tested for TRL 5, is tested in a relevant environment.
7. System prototype demonstration in a operational environment	Prototype near or at planned operational system.
8. Actual system completed and 'flight qualified' through test and demonstration	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents the end of true system development.
9. Actual system 'flight proven' through successful mission operations	Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation. In almost all cases, this is the end of the last "bug fixing" aspects of true system development.

In the following sections we discuss the state of the art of relevant research topics. We also discuss related government programs and policies that influence SI.

7.2.1.1 Knowledge Representation

The W3C developed a powerful expressive language, called Web Ontology Language (OWL), to express phenomenology as ontologies. Common Logic is a logic framework developed by ISO. It is intended for information exchange and transmission. The framework allows for a variety of different syntactic forms, called dialects, all expressible within a common XML-based syntax and all sharing a single semantics. Common Logic has some novel features, chief among them being a syntax which permits 'higher-order' constructions such as quantification over classes or relations while preserving a first-order model theory, and a semantics which allows theories to describe intentional entities such as classes or properties. It also fixes the meanings of a few conventions in widespread use, such as numerals to denote integers and quotation marks to denote character strings, and has provision for the use of data types for naming, importing, and transmitting content on the World Wide Web using XML. The following sections detail a list of necessary knowledge representation capabilities which we believe are lacking:

Spatial and Temporal Ontology

Tactical military operations occur in space and time. Accommodating spatial and temporal information is therefore a critical factor in successful M2M information transformation. The research should provide a coherent semantic model of space and time that can be shared within and between COIs. Space is an extension of a three-dimensional region in which entities of interest to a COI exist. For example, an answer to the question "where is the Eiffel Tower" to a person, who is few blocks away from its location, is different from answering the same question to a person who is in Washington, D.C. In the earlier case, the exact directions are required. In the latter case, however, it might be sufficient to say "Paris, France." Obviously, reasoning at the coarse space is different from reasoning in higher granule spaces. Existing spatial ontologies do not account for categories of space and their relationships. Therefore it is important to distinguish between categories of space to support information transformation in space and time.

Ontology Management

When ontologies are used in a distributed and dynamic environment like that being developed by COIs, technology must support several ontology-evolution tasks, ranging from data transformations to change visualization. Providing this support is difficult, as the current ontology formal models do not allow the enforcement of specific development procedures and lacks information about provenance and pedigree. Support for ontology evolution becomes extremely important in distributed development and the use of ontologies. As with all COIs, the number of ontologies is growing on the one hand and the semantics of data evolve and change on the other hand. While ontologies are useful in semantic reconciliation and are indeed necessary for practical and performance considerations, they do not guarantee in and of themselves correct classification of semantic conflicts, nor do they provide the capability to handle evolving semantics or a mechanism to support a dynamic reconciliation process. Successful applications of ontologies in such de-centralized and distributed COIs require substantial support for change management in ontologies and ontology evolution. Given an ontology O and its two versions, V_{old} and V_{new}, a complete support for change management in an ontology environment includes support for the following change management tasks:

- **Data Transformation:** When an ontology version V_{old} is changed to V_{new}, data described by V_{old} might need to be translated to bring it in line with V_{new}. For example, if we merge two concepts A and B from V_{old} into C in V_{new}, we must combine instances of A and B as well.

- **Ontology Update:** When we adapt a remote ontology to specific local needs, and the remote ontology changes, we must propagate the changes in the remote ontology to the adapted local ontology.

- **Consistent Reasoning:** Ontologies, being formal descriptions, are often used as logical theories. When ontology changes occur, we must analyze the changes to determine whether specific axioms that were valid in V_{old} are still valid in V_{new}. For example, it might be useful to know that a change does not affect the subsumption relationship between two concepts: if A U B is valid in V_{old} it is also valid in V_{new}. While a change in the logical theory will always affect reasoning *in general*, answers to specific queries may remain unchanged.

- **Verification and Approval:** Developers need to verify and approve ontology changes. This situation often occurs when several people are developing a centralized ontology, or when developers want to apply changes selectively. There must be a user interface that simplifies such verification and allows developers to accept or reject specific changes, enabling execution of some changes, and the rolling back of others.

- **Data Access:** If data exists that conforms to V_{old}, we must often access this data and interpret it correctly via V_{new}. That is, we should be able to retrieve all V_{old} data via queries expressed in terms of V_{new}. Furthermore, instances of concepts in V_{old} should equate with instances of equivalent concepts in V_{new}. This task is a very common one in the context of the Semantic Web, where ontologies describe pieces of data on the Web.

Ontology Reuse

To enable SI and modular information transformation, ontologies must be layered and structured in such a way as to minimize ontology redundancy and conflict, and maximize reuse. For this purpose, it is important to differentiate between upper level ontology and domain ontologies. A Domain-specific ontology models a specific domain, or part of the world. It represents the particular meanings of terms as they apply to that domain. As shown in Figure 5, upper ontology is a model of the common objects that are generally applicable across a wide range of domain ontologies. It contains core ontology concepts and relationships that are shared between COIs, so that an upper ontology for a COI can be a domain ontology for another. Ontologies can therefore be structured in a hierarchy that promotes reuse.

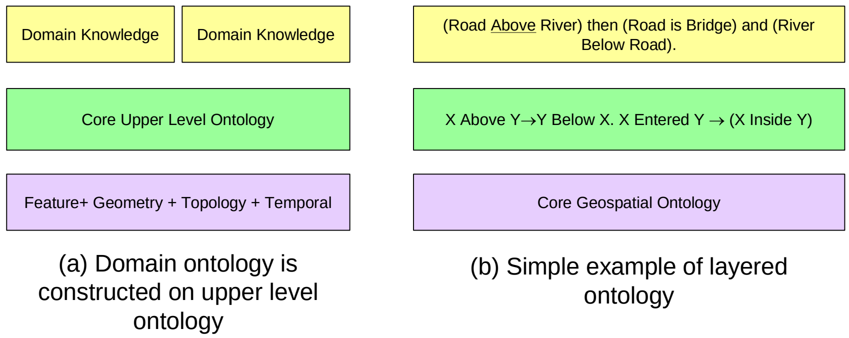

Figure 5, Layered Ontologies

7.2.1.2 S-O and O-O Mapping

Information Transformation within and between COIs is hard to achieve without knowing the semantic mappings between their ontologies. Manually finding such mappings is tedious, error-prone, and clearly impossible at the scale of multiple COIs. Hence, the development of tools to assist in the ontology mapping process is crucial to the success of SI. Several approaches have been proposed that employ machine learning techniques to find such mappings. Given two ontologies, the goal is to use a mathematical similarity measure that reliably finds the most semantically similar concept for each term in one ontology with that in the other ontology. It is therefore important to develop tools that provide multiple similarity measures that rely on machine learning (empirical feedback) to enhance match results.

The goal would be to map between services (input parameters and output) or databases (tables, columns and keys) and the corresponding concepts, relations and attributes of the ontology within the same COI.

7.2.1.3 Emergent Ontologies

Ontology is used to describe certain aspects of reality and a set of explicit assumptions regarding the intended meaning of the terms and relationships. Individual users may need to describe phenomenology using vocabulary that is not readily available in the domain ontology that is provided to them. We believe the same is true for most COIs. It is therefore anticipated that data sources and information services may come, go, and evolve over time. Ontology by definition encourages a top-down authoritative perspective and a rich model of controlled vocabulary. SI on the other hand encourages members within and between COI to share data with meaning. As meaning evolves, information transformation should accommodate this evolution so that new meanings can always be transformed and conveyed to interested parties within and between COIs. Therefore, viewing ontology as a rigid and controlled vocabulary will inhibit SI, in general, and information transformation in particular. Semantically-aware systems should support user-defined collaborative extensions to ontologies as the semantics of data sources change or evolve. Research in folksonomy should account for validation, management, and re-use of extended user defined ontologies.

7.2.1.4 Services

To date, the activity of creating Web processes using Web services has been handled mostly at the syntactic level. Current composition standards focus on building the processes based on the interface description of the participating services. The limitation of such a rigid approach is that it does not allow businesses to dynamically change partners and services. The ultimate goal set forth in this effort is to enable M2M interoperability so that COIs can semantically interoperate within and between information spaces such as those that ride on the GIG. The goal of M2M interoperability cannot be achieved without extending SI to also include the semantics interoperability of services and workflows within and between COIs. M2M can be enabled by software agents that should be able to discover, invoke, compose, and monitor resources offering particular services and having particular properties, and should be able to do so with a high degree of automation if desired. Powerful tools should be enabled by service descriptions, across the Web service lifecycle.

7.2.1.5 Context

Important to information transformation is gaining an understanding of the context of an information request. Is the statement: “All vehicles are armored” always true? Non-military vehicles are not armored. This statement, however, might be entirely true if its context was “military theater of operations.” Here we have added context to the statement, the context being “military theater of operations.” Several domains have already elaborated their own working definition of context. In human-machine interaction, a context is a set of information that could be used to define and interpret a situation in which agents interact. In the context-aware applications community, the context is composed of a set of information for characterizing the situation in which humans interact with applications and the immediate environment [Dey, 1998]. In artificial intelligence, the context is what does not intervene directly in problem solving but constrains it [Brézillon, 1999a]. Our working definition of context is that it is a collection of relevant conditions and surrounding facts that make a situation unique and comprehensible.

Simply stated, Context is the set of assertions and conditions under which a set of axioms are true. In a M2M interoperability setting systems should be able to indicate that a given a set of ontology axioms can only be determined within context. The basic relation relating context and ontology is *ist (C, O)* which asserts that Ontology *O* is true in Context *C*. In other words, ontology is only valid within a certain context. Context models are ontologies in their own right; we use OWL to create context models. This means that any ontology can play the role of context in a domain while being the ontology in another domain. The general agreement in the literature states that *ist (C, O)* in itself is true in a larger context, i.e., it occurs in a larger context.

Context defines metadata relating to its meaning, properties (such as its source, quality, and precision), and organization. More formal, context is the set of assertions and conditions under which a set of axioms are true. Ontologies are shared models of a domain that encode a view which is common to different communities. Context is a model that cast a local view of shared models, i.e., shared ontologies. Context can be considered as a filter that helps scope the subset of an ontology that is relevant to a given situation.

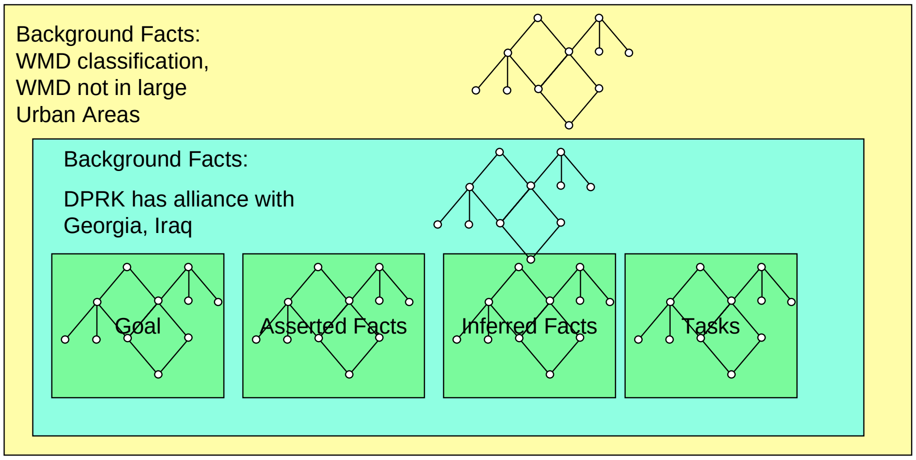

Figure 6, Context is background information about a set of domain facts

Using the Semantic Web terminology, context is a collection of metadata facts about a set of domain facts. As seen in Figure 6, both context facts and domain facts are graphs. To achieve this goal, a graph that represents domain knowledge must have a unique ID so that the graph that represents its context can reference it. This mechanism is known as Named Graphs. OWL is built on the graph model of the W3C Resource Description Framework (RDF) graph model. Current OWL specifications do not allow named graphs. Future versions of OWL are expected to support this capability.

7.2.1.6 Information Assurance

Security of information and services with considerations to space and time parameters are lacking. Security of information and services need to be dynamically determined in real-time depending on a user's situational context. Most importantly, it also needs to be determined based on the security implications from aggregating and inferring information from sources that are not sensitive.

It is important to distinguish between Authentication, Identifications, and Privacy. Authentication refers to the authentication or verification of a claimed identity. In other words, the user wishes to log on to a network or service, or undertake an online transaction and claims to be a certain person. The authentication process seeks to verify this claim via the provision of a characteristic (PIN, password, token, biometric, or other information), or multiple characteristics, known to be associated with the claimed identity. There is therefore a one-to-one matching process involved, as the characteristic in question is matched against the reference associated with the claimed identity, according to predefined threshold criteria in the case of biometrics.

Identification seeks to identify a user from within a population of possible users, according to a characteristic, or multiple characteristics, which can be reliably associated with a particular individual, without an identity being explicitly claimed by the user. There is therefore a one-to-many matching process involved against a database of relevant data. We should perhaps make a further distinction between identifying an individual from within a known population using relevant characteristics (PIN, password, token, biometric etc.) and seeking to identify an individual via connectivity address information. In the latter case, we may correctly identify an address and the name that is registered in association with it, but that does not necessarily guarantee that the same individual undertook a specific transaction (unless robust biometrics has been used across multiple processes). Privacy seeks to protect an actor's personal information, location information, access habits, and personal profiles from being accessed by unauthorized entities or services.

7.2.1.7 Enterprise Integration Tools

Enabling SI through the use of ontologies will provide substantial benefit to COIs whose data ride on the GIG. One side effect of this technology is that business logic will be encoded in ontologies and rules rather than being hardwired in the software application code. The picture gets even more complicated when ontologies and rules become part of the development life cycle of large enterprise solutions like the GIG. This essentially means that the management and development of ontologies and rules must be part of the overall systems engineering lifecycle tasks.

The goal of the Model Driven Architecture, defined by OMG, is to achieve maximum level of interoperability between large-scale distributed enterprise systems. Using the Model Driven Architecture methodology, system functionality is defined as a Platform Independent Model (PIM), using an appropriate modeling language and then translated to one or more Platform Specific Models (PSM) for the actual implementation. To accomplish this goal, the Model Driven Architecture defines an architecture that provides a set of guidelines for structuring specifications expressed as models. The translations between the PIM and PSMs are normally performed using automated tools. OMG recognizes the need to manage ontology development as an integral part of large-scale enterprise solutions. Led by IBM, a working group within OMG is currently developing an Ontology Definition Metamodel (ODM) as part of the Model Driven Architecture. ODM represents the foundation for an extremely important set of enabling capabilities for Model Driven Architecture-based software engineering, namely the formal grounding for representation, management, interoperability, and application of business semantics. The ODM specification offers a number of benefits to potential users, including:

- Options in the level of expressivity, complexity, and form available for designing and implementing conceptual models, ranging from familiar UML and ER methodologies to formal ontologies represented in description logics (OWL) or first order logic
- Grounding in formal logic, through standards-based, model-theoretic semantics for the knowledge representation languages supported, sufficient to enable reasoning engines to understand, validate, and apply ontologies developed using the ODM
- Profiles and mappings sufficient to support not only the exchange of models developed independently in various formalisms but to enable consistency checking and validation in ways that have not been feasible to date
- The basis for a family of specifications that marry MDA and Semantic Web technologies to support semantic web services, ontology and policy-based communications and interoperability, and declarative, policy-based applications in general.

7.2.1.8 Technology R&D Summary

The explosive growth of the Semantic Web in the commercial and academic spheres has already pushed the state of the art a considerable distance in a few short years. As more mainstream commercial interests begin to adopt semantic enablement for a myriad of business applications, this research is likely to increase in pace and emerge an ever more mature base of enabling tools and technologies. For these reasons, the gaps that this study identifies will likely close without wide scale government investment. The adoption of these technologies into the USAF operation setting, however will likely lag behind the commercial sector and thus will require government investment to prove its merit and trustworthiness and to acquire, adapt, and install in key programs.

7.2.2 Contribution of Existing DoD SI Programs and Activities

The FAA developed an architecture of functions that semantic technologies must perform to meet department requirements for interoperability. Significant progress has been made within the department to implement many of these functions. This section of the FNA examines the current state of implementation of SI technologies within the DoD with special emphasis on the interoperability-challenged strike mission area.

Our method for this survey took several forms. First, we participated in OSD NII's COI Forum and examined COIs that were active in the strike mission area (consonant with the study use case). This was useful in that it allowed us to study the state of implementation of network-centric data sharing policy. Second, we identified DISA's GIG and NCES programmatic activities that had as their purpose establishing facets of the GIG or which claimed to implement network-centric enterprise services. This was useful in that it assessed progress towards formal development of the shared services that DISA believes will enable the GIG environment. Third, we tracked down artifacts (e.g., data models, schemas, program briefs, and vocabularies) from the COI's and from exemplar programs which we believed were addressing facets of SI. This was useful in that it exposed early SI lessons learned and allowed us to assess the expressiveness and comprehensiveness of a number of 'state-of-the practice' schemas.

The survey was by no means comprehensive enough to cover all programs, projects, data model development efforts, and demonstrations, but we feel that it was comprehensive enough to determine the current level of SI implementation with sufficient breadth and depth to determine capability gaps for the FSA stage of this study. The "M," "C," and "D" sections of Appendix B list the resources we located and consulted for this part of the study.

7.2.2.1 Strike Mission Area Communities of Interest

Several COIs are working to build domain semantic models and services to support the strike mission area. We surveyed the following COIs: Time Sensitive Targeting (TST), sponsored by AFC2ISR and JFCOM/J8; Blue Force Tracking (BFT), sponsored by JFCOM/J8; Maritime Domain Awareness (the other MDA), Sponsored by the US Navy; Air Operations, sponsored by USAF ESC; Joint Targeting Intelligence, sponsored by JCS/J2T; Joint Track Management (JTM), sponsored by USN PWM-150; Joint Air and Missile Defense (JAMD), sponsored by JTAMDO; ISR, alternately sponsored by DIA, STRATCOM, and the Distributed Common Ground Station (DCGS) program; and Strike, sponsored by USSTRATCOM. Each of these COIs have seen varying degrees of resources and each have built semantic artifacts and/or services to capture and/or exchange domain knowledge. None of the COIs have followed an identical course of action as detailed policy for COI operations is still evolving. Similarly none of the above COIs have enjoyed a constant funding stream or a consistent command emphasis and sponsorship over time. The COIs usually feature participants from multiple services and organizations, most of whom are participating on a voluntary basis. In some cases, the "all are included" membership practice most COIs entertain has led to personality and/or service domination. Some, but not all of the COIs contain both SMEs and people familiar with data modeling and semantics. Those without the latter typically also lack the tools necessary to build executable schemas and ontologies.

Perhaps the most difficult thing that COIs must do is achieve consensus on key domain concepts, terms, and definitions. This act usually requires the participants to overcome service or command biases and to work towards what amount to semantic compromises. Since many of the

COIs are trying to define similar or overlapping mission area concepts - sometimes in a serial fashion and sometimes in parallel – and are intermittently aware of each other's activities, inheritance and reuse of terms and concepts can not always be realized. All the COIs are motivated to "get it right" and adequately represent their domain or sub-domain's equities. This eagerness has at times resulted in sub-optimizing behavior (invent all new terms quickly) or over-optimizing behavior (continually trying to achieve the total buy-in from all possible parties and/or forcing reuse of terms and constructs which do not properly capture domain concepts).

Viewed through time, one sees several interesting COI lifecycle phenomena. The COI lifecycle usually begins with great fanfare and several well attended organization meetings. Typically COIs then fragment into several working groups such as data modeling and vocabulary development. These working groups or panels meet over time and gradually dwindle to a core group of motivated (or resourced) individuals. The smaller groups typically coordinate the action of placing terms and their representations into a schema that is eventually published. If the COI is a pilot, it may also stage an interoperability demonstration to show the new schema at work. A recent trend is that as COIs become smaller and more focused (and thus less visible), other related organizations perceive the need for a new COI and subsume the work of the maturing COIs. These second generation COIs have the luxury (or burden depending on their point of view) of choosing between the work of earlier COIs or starting over. As a general trend, well-defined terms and vocabulary are carried forward and in general the sub-domain schema expands and improves. The strike mission area COIs have followed this pattern, the resulting winnowing effect is that the later COIs express schemas that are more universal and serve to bridge sub-domains with more ease.

Despite the many issues that the COIs have and continue to experience, most of them have embraced the challenge and at a minimum have developed schemas and/or controlled vocabularies. On the whole, the COIs have honored the department's policy and published their semantic products – for the most part structural metadata – with the DoD metadata repository. What has not happened yet, and will ultimately be necessary to enable M2M SI is for the COIs to begin constructing semantic ontologies for their respective domains and the mapping of correspondences between the domain and sub-domain ontologies. While this higher level of expressivity has always been part of the intent of the DoD's net-centric data strategy in general and the COI concept in particular, the enormity of achieving consensus on domain terms, establishing this knowledge as basic structural metadata, and governing the process of doing so in a resource and tool poor environment has so far not realized this higher goal. The output of the COIs none-the-less are an important first step toward SI.

During 2006 and extending into the present, the "big bang" which originated many COIs in the 2004-05 timeframe has seen certain mission areas begin to form aggregate COIs or COI portfolios for certain key mission areas. In the strike operations mission area, STRATCOM initially proposed a Global Strike COI which eventually became simply the Strike COI. This well resourced COI took as its input the work (and many of the participants) of the TST, BFT, MDA, JAMD, Air Ops, and the JTM COIs. Strike has since attempted to define a common "middle" ontology to link the concepts in these sub-communities into a composite schema for global strike operations. To this end, the Strike COI is trying to decide which schema to adopt as 'core' schema. The two leading candidates at present are the Army and international combined armed community's Joint C3 Integrated Exchange Data Model (JC3IEDM) and the Joint Track Management (JTM) schema. The Strike COI is also investigating using the "tearline" concept successfully demonstrated by the Cursor-on-Target program (described below) which involves a simple, common tag set for all of its objects with a means to attach domain or program detail to the common header.

At a higher organizational level than the aggregate COIs, JFCOM J87 has established what it is calling the C2 Portfolio to bring governance to the scattered community of COIs that are working Joint Task Force-level C2 interoperability. This body, which is being stood up in early 2007, is selecting COIs willing to follow joint portfolio governance and that can contribute to the joint C2 interoperability challenge. In exchange for this cooperation and to in part to realize it, the C2 Portfolio is investigating how to resource COIs with common tools, expertise, and guidance. We fully expect that the C2 Portfolio will also push for the development or adoption of some common or core group of terms to bridge and relate the constituent COI schema efforts.

7.2.2.2 Core Schema and Upper Ontology Efforts

The concept of a common core schema is also being explored by OSD NII's Core Data Model Working Group as a possible means to relate many COI data models to each other. Certain key schematic elements, notably the Open Geospatial Consortium's Geography Markup Language (GML), the Intelligence Community's Metadata Standard for Information Security Marking (IC-ISM), and the DoD Discovery Metadata Standard (DDMS) have all received popular support as core standards. These particular schemas have also been promulgated in recent executive guidance such as EO 13338 (P24). Several programs are also trying to evolve "upper" ontologies as means to establish high-level concepts suitable to ground high detail domain schemas developed by the COIs to a stable upper structure of concepts. A recent multi-department community of practice (COP) has been formed to look into this concept. Similarly, the Defense Intelligence Agency has begun a project titled the "DODIIS Upper Ontology" in part to develop a common high level concept space for intelligence analysis.

7.2.2.3 Recent Strike Domain Programs with a Semantic Component

Prior to the COI movement described above, JFCOM's Family of Interoperable Operating Pictures (FIOP) program in 2002 launched an effort called the FIOP Network Based Services (NBS). NBS had as its goal a capability called Cross Service Weapon Target Pairing (XSWTP) which succeeded in establishing mappings between the schemas of the principal fire control systems of the joint services (eventually about eight programs in total). The method to this end was fairly labor-intensive and resulted in a massive Excel file that defined the many-to-many maps between the permutations of the system schemas. This effort was defunded before it reached its demonstration objectives but was a bellwether to the community of the complexity of deterministically working between sub-domains. During this timeframe, a second program developed at ESC called Cursor-on-Target approached the problem in a different fashion. CoT developed a very simple schema that in nine to ten elements expressed the basic target information details that most systems actually wanted and needed for situation awareness. Tacked on this simple envelope was a means for programs to add detailed fire control or C2 information that they wished to share with other programs that needed this level of complexity. CoT was boosted to community-wide attention in 2004 when some 300 programs participating in Joint Expeditionary Forces Experiment (JEFX04) were successfully adapted in a matter of weeks to exchange CoT envelopes. JEFX04 measured millions of CoT exchanges.

Several strong data models have emerged since the early 2000s that have benefited greatly from the FIOP-NBS and CoT lessons and the governance and schema-winnowing activities of the COIs. These include (among quite a few others) JC3IEDM (Army), the Common Battlespace Object (JFCOM), the Common Mission Definition (USAF ESC), the Operational Joint Architecture Working Group Data Model (OJAWG), and the Joint Track Management Data Model (PEO C4I). Documentation for these models is noted in the "D" section of Appendix B.

We expect to see these models (which at this time are largely implemented as well-developed schemas and not as ontologies) to continue to converge and improve as collectively the best practices and concepts are carried forward from each. We also believe that as the dominant models become stable and continue to see use they will eventually be enriched with more expressive semantics.

7.2.2.4 Network Centric Enterprise Services

DISA's NCES Program is directly addressing the infrastructural needs of the emerging GIG. In their own words, NCES offers "global information advertising and delivery services" for those who have information and "global services to find and retrieve information" for those who need it (see M19). NCES attempts to implement the network centric data strategy using a combination of services, interface specifications, and software development kits (SDKs). NCES' Increment 1 contains three product lines: Enterprise Collaboration (enabling the warfighter with collaborative tools including chat and conferencing tools), Enterprise Portal (common resource, information, and tools web space) and Content Discovery & Delivery (CD&D) capabilities. This last group features Federated Search (providing services to find and aggregate information across GIG enterprise data sources); Enterprise Catalog (providing services to store and retrieve metadata published from GIG enterprise data sources); and an Enterprise Content Delivery Network (providing services to store, cache, and forward-stage information for fast access). Underpinning these three product lines are what DISA refers to as its Service Oriented Architecture Foundation (SOAF) Capabilities. These include:

- Service Security - Providing services to create, manage and enforce access control policies for GIG enterprise services
- Enterprise Service Management: Providing services to monitor and manage GIG enterprise services and reporting of Service Level Agreement/QoS information to GIG enterprise service consumers
- Service Discovery: Providing services to publish and find GIG enterprise services registered and categorized in an enterprise registry
- M2M Messaging: Providing services to support reliable information exchange at the machine level
- Mediation: Providing services to support translation between different message formats and the creation and implementation of process workflows across the enterprise
- Metadata Registry: Providing an ability for enterprise systems to discover, store, and manage various metadata artifacts while promoting data visibility and reusability

The current status of the CD&D and SOAF elements of NCES is that they are presently in early prototyping stages prior to an industry down-select. Reference implementations of these capabilities exist within the various DoD networks/security enclaves. DISA's strategy appears to stress open industry standards and DISA has stated its desire for commercial solutions for Milestone B. Several COIs, notably the Command and Control Space Situation Awareness (C2 SSA) and Maritime Domain Awareness (MDA) COIs have executed pilot interoperability demonstrations that exercise the nascent CD&D and various aspects of the current SOAF implementation.

A review of the programmatic documentation for the CD&D and SOAF sub-systems (see "M" series documents in Appendix B) reveals that at this time their services are focused on structural vs. semantic metadata. It stands to reason however, that as the NCES begins to provide ever more capable structural metadata services and begins to see the buy-in of larger programs, it will begin to consider adding semantic metadata capabilities and services. It is logical to conclude that as the technology to provide these services is developed in the commercial sector, DISA's standards and commercial source acquisition approach to NCES could speed their availability to the DoD user community.

DISA also runs the DoD Metadata Repository (MDR) which can also been seen as a semantic program of note. As of November 2006, the MDR had 7,600 registered government users and featured 186,000 XML items (e.g., schemas and stylesheets). The MDR also offers translation logic for mediation and taxonomies. At this time, however, the MDR only provides services for structural metadata. It is likely that as semantic metadata become available, that the MDR will form a repository for it and begin to provide many of the services it now provides for structural metadata.

7.2.2.5 Program and Activities Summary

Our survey of DoD programs and activities, many the direct result of recent net-centric data policy, confirms that the department has made a good start towards semantic enablement. Its COI processes have made major steps towards the development of domain schema, services, and data models as well as raising stakeholder awareness of the need to assemble and manage these resources. Few of these efforts have progressed beyond structural metadata at this time however. Similarly, early GIG service implementations are beginning to provide structural metadata services to the broader community but also have not yet progressed towards providing semantic metadata and services. As such, DoD's current programs and activities are not yet providing the rich functionality described in the FAA's objective functional architecture but have the potential to do so as the technology becomes available.

7.3 SI Material Capability Gaps

Figure 7 summarizes our estimation of the relative maturity of the component semantic technologies (material solutions) described in Section 7.2.1. Figure 7 uses the Technology Readiness Maturity Levels (TRLs) shown in Table 1. It is clear that there are substantial gaps in almost all the components of SI. As Figure 7 indicates, Security, and Migration are the least mature capabilities, while some aspects of Domain Knowledge, Publish and Discover, and Tools are relatively mature. None of the surveyed technologies are operationally mature at this time however.

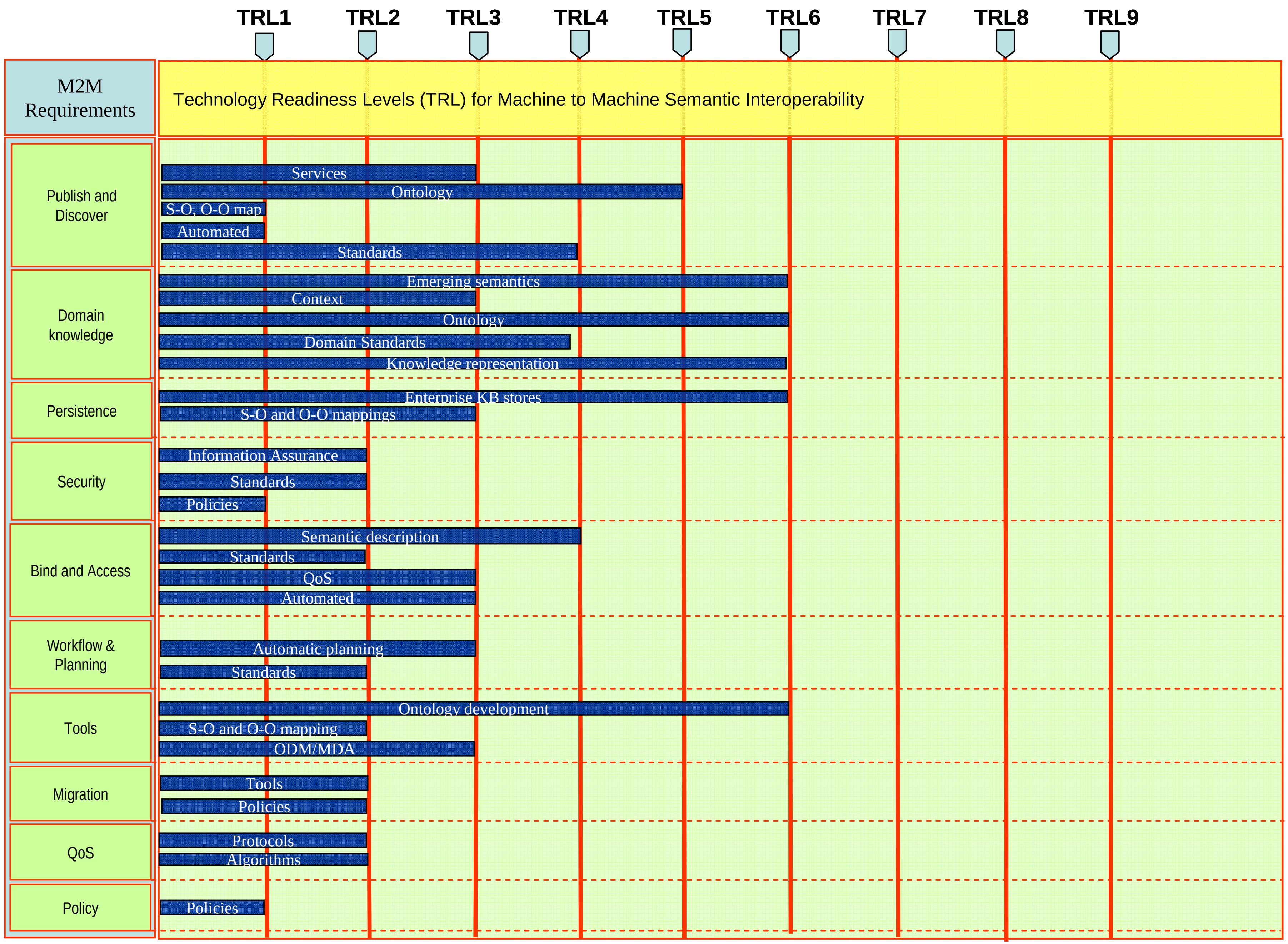

Figure 7, SI Technology Readiness Levels

In the following sections we will provide a breakdown of these gaps.

7.3.1 Domain Knowledge

Although OWL adds considerable expressive power to the Semantic Web, it does have expressive limitations, particularly with respect to what can be said about properties. Many of the limitations of OWL stem from the fact that, while the language includes a relatively rich set of class constructors, the language provided for talking about properties is much weaker. In particular, there is no composition constructor, so it is impossible to capture relationships between a composite property and another (possibly composite) property. We believe that OWL is not sufficient to capture many aspects of real world knowledge. The two examples below show use of domain knowledge for a USAF problem that cannot yet be represented in OWL.

If a Country X SA-2 battery has a single SPOON REST **and** three light-up along the border 30 km apart, **then** there are three batteries.
If the 5th Flying Wing is the only wing flying the L-39 **and** the L-39 is known to be associated with bio warfare training **then** the 5th Flying Wing is associated with bio warfare.

It is therefore important to extend OWL with Horn clause rules extension. This will add much needed flexibility to capture some of the behavioural aspects of the phenomenology.

To enable reasoning about actual real world objects and events, we must provide a way to map (subscribe) data to ontologies. The prevailing paradigm in information transformation and sharing is focused on exploring ways to associate ontologies to data elements and use hardwired ontology mappings to perform information transformation. Innovative technologies to support the development and operation of dynamic functional mapping between ontologies and information sources are required. The W3C is currently working on developing rule language that can be combined with OWL to provide a rich knowledge base. Finally, OWL standards use XML Schema data types but do not support derived or user defined data types.

Regarding Common Logic, we are not aware of any significant implementation of standards to support it. Standards are also important at the domain level. There is a clear evidence of inconsistent use of knowledge representation across COIs. Some use UML, some XML, and few use plain textual definition of vocabularies. An overall DoD enterprise ontology framework is needed. This framework should encourage re-use, standardized knowledge representation (e.g., OWL or Common Logic) and ontology lifecycle management.

A simplified view of the Semantic Web is a collection of Web-retrievable OWL documents, each containing only one graph. OWL enables us to describe any one graph and merge a set of graphs into one larger graph. However, it does not provide suitable mechanisms for talking about graphs using named graphs. The ability to express metadata (context) about graphs is required. Reasoning about facts in context poses many challenges in terms of the monoticity of the inferences. Such challenges must be resolved so that context can be embedded in knowledge models.

Present work environments lack the applications and the support tools to model and exploit context. The context associated with everything that impinges on the privacy of the GIG users with M2M capabilities, including the state and nature of available resources, the activities and status of GIG participants, and the context associated with their current status, potentially affects all actions and decisions about access control. Sometimes context is wired into data or into software processes. The problem is that the majority of ontology alignment efforts focus on mapping between concepts and relations. The ontology alignment problem can be described as follows: given two ontologies, find the relationships (e.g., equivalence or subsumption) holding between these entities. We observe two main problems with the prevailing approaches to ontology alignment:

- Alignment focuses on mapping between concepts of the underlying ontologies and ignores how instances in these ontologies are related. For an example, consider a populated ontology of cars and another of car accidents, alignment only between concepts would not allow us to know which car in the first ontology matches which accident in the second.
- Alignment does not consider the fact that mapping is not only restricted to equivalents or subsumption, but it should also include functional relationships at both type and instance levels.

In addition to the importance of ontology alignment in information transformation, it is also important to provide the means to logically reuse ontologies. The Web Ontology Language (OWL) defines the owl:imports construct, which allows one to include by reference all axioms contained in another knowledge base (KB) on the Web. This certainly provides some syntactic modularity, but not a logical modularity. Jim Hendler’s group at the University of Maryland pioneered the “E-Connections” approach as a suitable formalism for combining KBs and to eventually achieve interoperability between ontologies. We believe this approach to be beneficial and promising.

7.3.2 Publish and Discover

To enable Service Oriented Architecture, it is necessary to have technologies to support cataloguing and publishing, and protocols to support discovery of ontologies, S-O, O-O, and services. During discovery of services and digital content, it is important to consider not only functionality, but also the QoS of the corresponding activities. No commercial technologies are currently available to support these requirements.

7.3.3 Tools

Existing tools to support many functional requirements are lacking. Protégé, SWOOP and Altova Semanticworks are the most used tools for ontology development. The first two tools are produced by open source and do not provide necessary enterprise level maintenance and support. Semanticworks supports syntactic validation of the ontology; however, it does not support semantic validation.

Tools are also needed to support semi-automated and automated Ontology-Ontology and Schema-Ontology mapping. Furthermore, after extensive research, we are not aware of tools that support ODM in a collaborative environment as part of MDA life cycle. It is also essential to provide tools to support the migration process from legacy sources to semantic data stores. No such tools are currently available.

7.3.4 Persistence

Several systems currently exist for the storage of Semantic data, specifically RDF. These databases, often referred to as RDF stores, exist as both Open Source projects and commercial product offerings. Since Guha's first RDF store, several other Open Source stores have been developed. These include: Sesame, Threestore, developed, Redland, and Oracle 10G. Performance, scalability, security, and transactional capability remain major obstacles.

7.3.5 Migration

Migration, as defined by its functional requirements, is similar to the well known ETL for enterprise integration. ETL, stands for Extraction, Transformation and Loading. The Extract phase involves extracting and consolidating data from different sources. Each system may use different data formats, structures, and semantics. Common data source formats are relational databases and flat files. The Transform phase applies a series of rules or functions to extracted data to derive the data to be loaded. This usually involves developing an understanding of the semantics of both the source and target semantic data stores. Finally, the Load phase loads the data into the new semantic data store.

Within the context of SI, it is necessary to provide ETL technology to migrate from legacy sources to semantic based sources (e.g., RDF data stores). In other cases, it is more practical to leave legacy databases as is, and provide a virtual runtime semantic view from which they can be accessed and queried. Currently not commercial tools are available to support this capability.

7.3.6 Security

Semantic technology can play an important role defining the operational semantics of security protocols. A security protocol describes a number of behaviors. Each such behavior called roles. The OASIS organization developed SAML as an XML-based framework for communicating user authentication, entitlement, and attribute information. SAML allows business entities to make assertions regarding the identity, attributes, and entitlements of a subject (an entity that is often a human user) to other entities, such as a partner company or another enterprise application. SAML, however, does not allow reasoning about user roles and releasability. Ontology based security framework is still required. We are not aware of a substantial body of research that uses semantics for security management.

7.3.7 Quality of Service

To enable adequate QoS management, research is required to develop mechanisms that specify, compute, monitor, and control the QoS of the products or services to be delivered. The composition of workflows to model e-service applications differs from the design of traditional workflows due to the number of Web services available during the composition process and to their heterogeneity. Two main problems need to be solved: how to efficiently discover Web services and how to facilitate their interoperability.

To enhance workflow management systems with QoS management, it is important that a QoS model allows for the description of nonfunctional aspects of workflow components from a quality of service perspective. [Amit Sheth, 2006 Wfms] proposed an automated processes to compute the overall QoS of a workflow. The model is layered on OWL-S. The model is supported by a mathematical model called SWR. The proposed QoS model and mathematical model have been validated with the deployment and execution of a set of production workflows in the area of genetics. The analysis of the collected data proves that their models provide a suitable framework for estimating, predicting, and analyzing the QoS of production workflows. No commercial technology that supports semantically-based QoS is currently available.

7.3.8 Workflow and Planning

One means to approach the functional requirements for workflow and planning would be to enhance the current Web process composition techniques by using Semantic Process Templates to capture the semantic requirements of processes. Semantic process templates can be either based on OWL-S or WSMO and it can act as configurable modules for common processes maintaining the semantics of the participating activities, control flow, intermediate calculations, and conditional branches. The templates are instantiated to form executable processes according to the semantics of the activities in the templates. The use of ontologies in template definition allows a much richer description of activity requirements and a more effective way of locating services to carry out the activities in the executable Web process.

7.3.9 Binding and Access

OWL-S (formerly DAML-S) is an ontology of services that make binding and access functionalities possible. The overall structure of the ontology has three main parts: the service *profile* for advertising and discovering services; the *process model*, which gives a detailed description of a service's operation; and the *grounding*, which provides details on how to interoperate with a service, via messages. This will enable, for example, collaborative agents to perform tasks that require access to data and services within and between COIs and hence achieve maximum SI. The Web Service Modeling Ontology (WSMO) is another approach that has some similarities with OWL-S. The WWW Consortium SW Group is considering both these standards with the goal of developing a technology to enable conceptualizing and organizing semantic information about services. There are no commercial products that support OWL-S or WSMO at this time.

7.4 SI Non-Material Gaps

Section 7.3 describes gaps in semantic technology and/or its realization as actual tools. This section assesses non-material (DOT_LPF) gaps.

7.4.1 Doctrine and Policy

Figure 7 alludes to the fact there are also policy gaps with respect to SI. While the top-level department policy is completely consonant with SI, there is a profound lack of normative guidance as to how specifically it could be fully achieved. All too often, department policies and guidance make no distinction between structural and semantic metadata. It is hardly surprising then that programs with limited budgets and resources stop at the establishment of structural metadata and believe they have implemented the intent of the guidance. The existing guidance is sufficient, however, to drive the collection of domain vocabularies and schema but it must be made to also direct the development of domain ontologies.

7.4.2 Organization

At present, there exist at OSD and DISA organizations dedicated to the broader aims of network-centricity. OSD/NII has taken a very active role in implementing the department’s data sharing strategy by supporting and encouraging the COI movement. The COIs themselves have formed many quasi organizations to leverage domain experience and to build controlled vocabularies. The emergence of composite COIs and the JFCOM C2 Portfolio Manager suggest that COI governance is adapting its organization model to be more effective. On the whole, however, these organizations have had their hands full with the construction and management of structural metadata. Collectively, these organizations are probably the best structure to marshal the development of the domain ontologies needed for M2M SI. For this reason, we do not assess there to be significant organizational gaps.

7.4.3 Training

Semantic concept/technology training, on the other hand, is fairly immature. The cause of this gap is likely twofold, first those who need the training (largely the COI participants) are fully engaged in the tasks of vocabulary socialization, structural metadata compilation, and COI politics, and second, semantic technology is fairly immature and thus there isn't much to be trained on. The materials used by the COIs to self-train expose participants to the concepts of semantic ontologies but do not at this time train individuals how to compile and manage them. Similarly, there is not much in the way of SI lessons learned or best practices to serve as learning aids or references. Consequently, we see the lack of training opportunities and materials as a gap.

7.4.4 Leadership

The need to prepare service programs and activities to operate within the GIG has been promulgated from the highest echelons. Furthermore, considerable attention has been paid to establishing Chief Information Officers and information managers at various levels of the DoD and USAF. We do not perceive that there is a leadership gap with respect to encouraging the development of interoperability technologies. We do perceive however, a growing expectation that all the attention paid to net-centric enablement will soon payoff in big ways. In some cases, there is also frustration that the investment made in COIs has not fully solved key interoperability challenges. If leadership is not eventually shown that SI is viable and effective, it may become more difficult to obtain enabling resources. At the present, however, we do not consider leadership to be a considerable gap.

7.4.5 Personnel

Presently, the personnel who are leading the charge towards SI are a mixture of domain savvy people and technologists. This is probably the right mix to get SI technology explored, developed, and demonstrated. A gap that is evident however, concerns the availability of knowledge engineers. Knowledge engineers ultimately are going to be needed first to assist with the orderly compilation and expression of domain knowledge, then later in the maintenance and operation of the knowledge-enabled enterprise. While this expertise may be found in academia and within certain contractor staffs, the USAF has neither a knowledge engineer AFSC, nor the means in place to train its own knowledge engineers. This gap will become more evident as SI technology begins to transition to operational practice.

7.4.6 Facilities

Facilities – in the form of networked resources – now exist to store and index structural metadata on the unclassified, SIPRNET, and JWICS LANs. SI technology is going to need similar provision for semantic metadata. At present, this is a gap.

7.5 FNA Summary

It is clear from the FNA gap analysis that the majority of SI material technologies fall between TRL 1 and TRL 4. Technologies that relate to ontology development, reasoning, emerging semantics, and knowledge representation are at a stage where they can gradually be transitioned to TRL 6, TRL 7 and TRL 8. The others will need greater attention before they become viable. Notwithstanding, the uptake of the technology must be supported by appropriate policies, organizations, leadership, facilities, and trained personnel. We have shown that most of these enablers have implementation gaps too. Across the DOTMLP, we discovered no barriers to the implementation of SI technologies, nor did we see any evidence of wasteful overlap or overmatched technology.

8 Functional Solutions Analysis

8.1 Method

The final stage of this study is the FSA which examines and/or recommends potential DOTMLPF and policy approaches to solving (or mitigating) the capability SI gaps identified in the FNA. The FSA stage further recommends an approach to developing a future research and development agenda for achieving SI within and between USAF systems, with M2M interoperability as the ultimate goal. This agenda is in the form of a roadmap that identifies future directions and topics for research and development, as well as goals, objectives, needs, priorities, and risks. The FSA stage also defines the recommended methods and procedures to achieving the plan in terms of short-term studies, long-term studies, prototyping activities, operational tests, and so on, with estimated budget, major tasks, milestones, and schedules. The completed FSA documents the capability gaps and recommended research approaches and includes traceability back to the originating capability requirements and issues. Finally, the FSA also highlights ways other than material solutions that the USAF might address SI capability gaps (such as leveraging the COI governance model or promulgating key standards).

8.2 Findings

Our challenge in this section was to recommend a systematic approach to filling the DOTMLPF gaps presented in the FNA. First we will address the material solutions as they represent the bulk of the gaps.

8.2.1 SI Material Solution Roadmap

Figure 8 shows our estimate of the probable maturity path of the semantic technology gaps exposed by the FNA. This estimate is based on what we believe the commercial sector is likely to develop and deliver over the next ten years. In the figure, we show the estimated placement of key TRL milestones (the diamonds), technologies (dark bars), and technology manifestations including tools, SDKs, and engines (light bars). From this estimate, it is possible to speculate that SI and supporting technology will reach FOC within the commercial sector within the next three to six years. Assuming that the government will continue to acquire the bulk of its IT infrastructure from the commercial sector, it is also possible to speculate when these technologies will naturally enter the government sphere through standard acquisition channels after it has been regularized in the commercial sector. In this view, SI technology could enter regularized government service beginning in the FY10-11 timeframe assuming that the investment has been made to develop domain ontologies in advance of its arrival.

This view, however, discounts the possibility of the government remaining an actor in the development and proof of SI technologies. Historically, the government through defense and space program investment has been the genesis of many commercially successful technologies and is often an early adopter of nascent attempts to commercialize these technologies. Given the track record of the private sector to rapidly mature and productize basic technology components, the optimal roles for federal defense investment in an emerging technology would be to support the following activity categories:

- Invest in the identification and proving of promising theories and approaches
- Accelerate the maturity of emerging technologies that solve critical defense needs
- Prove that an emerging or maturing commercial technology can improve department functions, offer cost savings, and/or reduce risks
- Make ready for the adoption of a disruptive technology

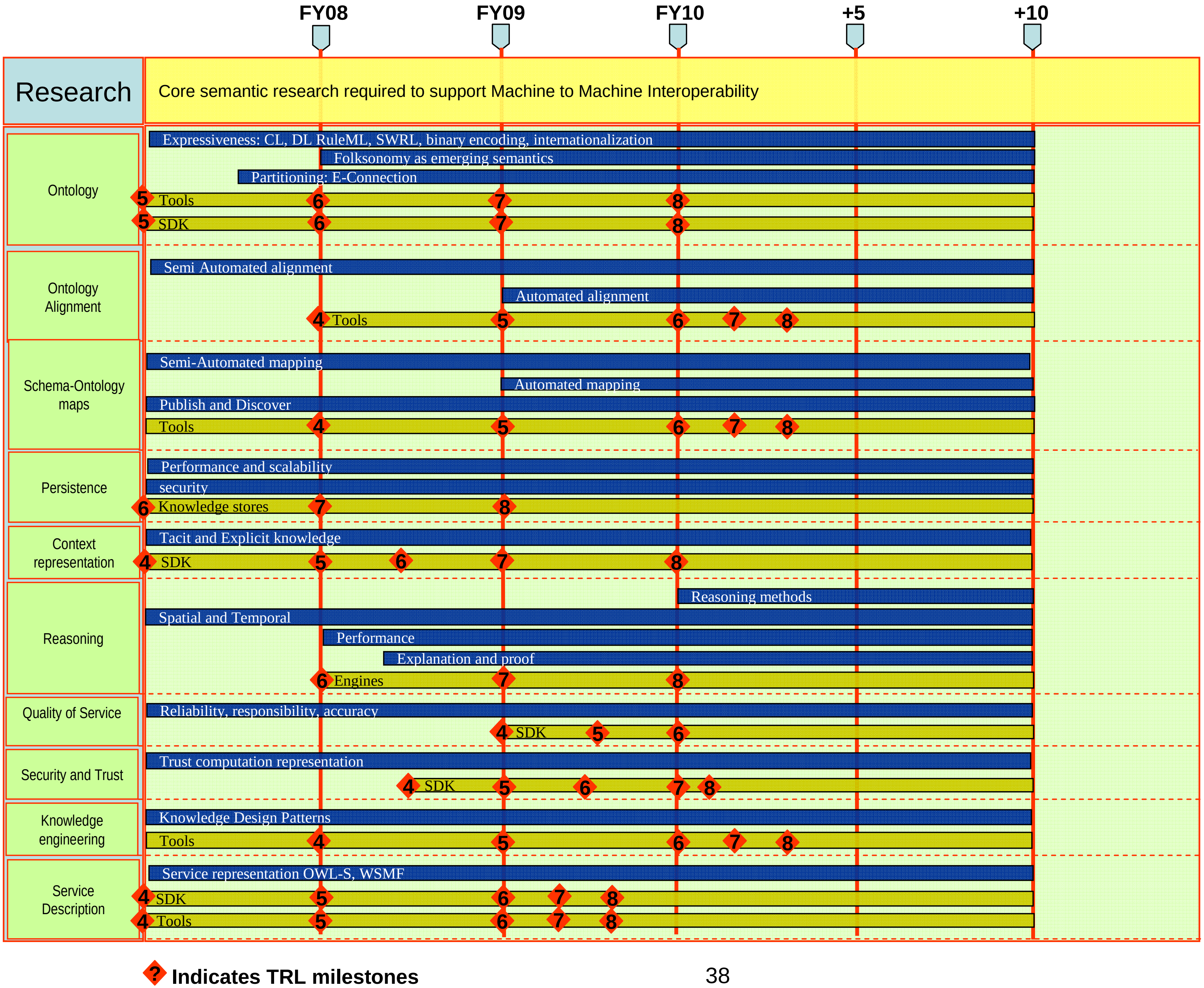

Figure 8 TRL milestones overlaid on research topics to support SI development

Presuming the government does wish to invest its own resources, the question of what to invest in and in what sequence should be modulated by the government's priority with respect to the four activities above and the availability and color of investment funds. The theme that runs through the four activities is a graduated degree of technology maturity, this helps to select technologies that are germane to any one of the activities.

To assist with the determination of a possible investment priority scheme, it is also possible to approach SI technology development from a phased maturity perspective that is agnostic to priorities. Even without priorities, such a perspective is useful as it prescribes specific activities that must be performed given that one has chosen to be an exponent in the maturity of a particular technology. In this view, it is possible to suggest activities to further mature technologies at any particular TRL. Table 2 is a regeneration of Table 1 with a column added to suggest maturing activity that should be performed to raise a particular technology to the next maturity level and specific goals associated with that TRL change.

Table 2, Technology Readiness Levels with maturing activity and goals

Technology Readiness Level	Description	Maturing Activity / Goals
1. Basic principles observed and reported	Scientific research begins to be translated into applied research and development.	Develop and demonstrate basic principles as a simple application via proof of concept approach.
2. Technology concept	Invention begins. Once basic principles are observed, practical applications can be invented.	Develop and demonstrate basic principles via proof of concept approach, explore limitations and assumptions.
3. Analytical and experimental critical function	Active research and development is initiated.	Develop and demonstrate via proof of concept approach, explore interaction with other technology components.
4. Component validation in a laboratory environment	Basic technological components are integrated to establish that the pieces will work together. This is relatively "low fidelity" compared to the eventual system.	Develop and demonstrate via proof of concept approach that the technology components work together and could support a relevant environment.
5. Component validation in a relevant environnent	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements.	Configure technology components into a functional system prototype that could support a relevant environment. Demonstrate and verify design.
6. System/subsystem prototype demonstration in a relevant environment	Representative model or prototype system, which is well beyond the breadboard tested for TRL 5, is tested in a relevant environment.	Develop and demonstrate system prototype in a favorable operational environment, measuring how well technology meets requirements / user needs.
7. System prototype demonstration in a operational environment	Prototype near or at planned operational system.	Mature prototype into stable form focusing on scalability, reliability, supportability, and improved performance. Demonstrate in typical operational environment.
8. Actual system completed and 'flight qualified' through test and demonstration	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents the end of true system development.	Continue to focus on scalability, reliability, supportability, improved performance, fault tolerance, and deployability. Demonstrate in stressing operational environment.
9. Actual system 'flight proven' through	Actual application of the technology in its final form and	Monitor system performance, execute preplanned product improvement plan.

successful mission operations	under mission conditions, such as those encountered in operational test and evaluation. In almost all cases, this is the end of the last "bug fixing" aspects of true system development.	

Applying the maturing activities to any particular SI technology would presumably advance it to the next level. A generic approach would recommend conducting the appropriate maturing activities to all SI technologies until they are all operationalized. This approach however is insensitive to the ongoing work of the commercial sector, and the current posture of the SI technology portfolio within the USAF.

In Section 7.4.4 of the FNA, we identified the issue that USAF leadership presently has high expectations for SI and that there also exists some frustration that the COI method of governance has yet to deliver big returns on the investment made in it to date. Similarly, the authors have experienced DoD/IC stakeholder perceptions that semantic technology is immature and experimental and at best a boutique technology. It is the considered opinion of the authors that to continue to enjoy favorable sponsorship from USAF leadership and to raise the overall community awareness of SI technology, SI technology is going to have to be demonstrated as an effective interoperability solution that is practical in the air operations arena.

Consequently, our recommended roadmap approach places early emphasis on the demonstration of TRL 6-8 capabilities that are ready for transition and uses these as the catalyst for executing and updating the other less mature technologies. Figure 9 depicts this notion in a lifecycle perspective. The six steps in this "SI Roadmap Lifecycle" accord the highest priority to the step at the 10 o'clock position "Develop and demonstrate TRL6." We feel that this step and the two that follow it which project TRL7-9 technologies into operational settings are needed to provide SI technologies an early and convincing win within the USAF.

A technology investment roadmap would result when the sequential and numbered steps shown in Figure 9 are systematically applied to the maturing technologies shown in Figure 7 based on their present TRL. This approach however would not address the amount of time needed to mature each of these technologies. To gain this perspective, it becomes necessary to place them into short term and long term categories.

8.2.1.1 Short Term Investment Strategy

In the short-term we propose investment in demonstration capabilities that are possible to conduct in segments of six to twelve months in duration. The objectives of the short-term strategy would be:

- Generate interest within the USAF by demonstrating TRL6-8 technologies leveraging the demonstration use cases related to USAF GIG programs (the strike use case family)
- Demonstrate the feasibility and applicability of SI technologies in solving M2M interoperability issues
- Empirically demonstrate the maturity of current semantic technologies
- Determine the technology components that can be transitioned to operational pilots, i.e., TRL 7
- Determine technology components that require further R&D before being transitioned
- Determine specific policy aspects that are necessary to speed the transition and wide adoption, i.e., TRL 8, of the new technology

- Update the functional requirements resulting from the execution of the short-term plan and further revise the investment strategy

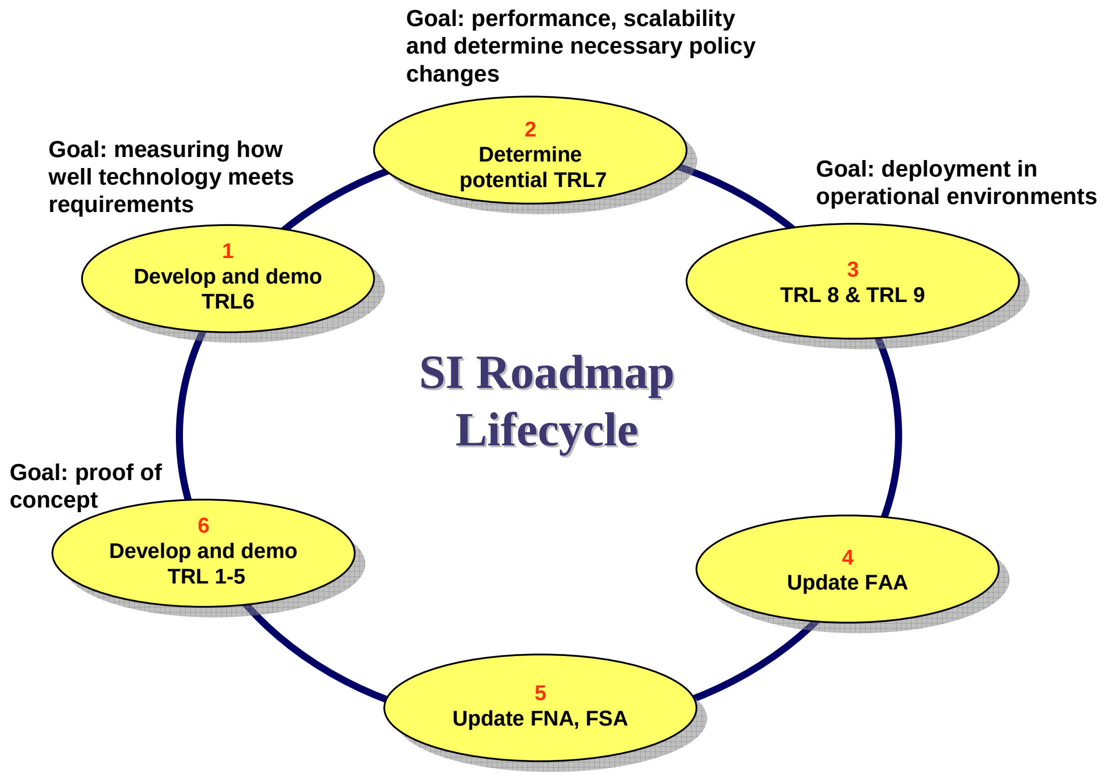

Figure 9, Lifecycle components of SI roadmap

Reordering the four investment categories with respect to the short term reality, places them in this priority order:

1. Prove that emerging and maturing commercial SI technology can improve department functions, offer cost savings, and/or reduce risks
2. Accelerate the maturity of emerging SI technologies that solve critical defense needs
3. Invest in the identification and proving of promising SI theories and approaches
4. Make ready for the adoption of disruptive SI technology

Integrating these relative priorities against filling FY07/08 gaps from the FNA would suggest the following choice of activity targets for potential short term investment:

- Prove via demonstration activities:
  - SI use of Domain Knowledge via Ontology (Expressiveness, Folksonomy, Partitioning)
  - Effect of Persistence (Performance and Scalability)
  - Spatial/temporal reasoning
- Accelerate via contracted effort capabilities to develop:
  - Ontology Alignment
  - Schema-Ontology Maps
  - Knowledge engineering tools
  - Service description

- Invest in basic research toward approaches towards:
    - Quality of service
    - Security and trust
- Consider infrastructural investment in:
    - Ontological repository services in advance of NCES
    - Tools to assist COIs

8.2.1.2 Long Term Investment Strategy

The scope of the long-term plan is to invest in R&D to advance technologies that are between TRL 1 through TRL 5 if the commercial sector has not already done so. An exception to this is technologies and tools that would assist COIs with domain knowledge development. We believe that this basic yet, immature technology application should see early investment as it is critical to the ultimate success of the COI contribution to future SI success. We therefore recommend that the government starts the long-term plan shortly after concluding the first cycle of demonstration development in the short-term plan. This will enable the government to generate interest in the community by realizing the immediate benefits from semantic technologies and hence pave the way for COI adoption of ontology to develop their domain knowledge.

The objectives of the long-term plan include:

- Encourage COIs to build their respective domain ontologies. Designing domain ontologies by committees is a long process. Starting such a process early will enable the government to use these ontologies a few years later when deploying semantically interoperable solutions. To following sub-objectives are important:
    - Utilize available ontology tools
    - Standardize the ontology encoding standards (e.g., OWL)
    - Provide online and secured catalogues to publish and discover ontologies
    - Provide guidelines and best practices to develop domain ontologies
- Provide sustained and gradual technology insertion to the SI technology evolution
- Advance and speed the development of high priority lower TRLs (TRL1 through TRL5). Those priorities are determined by prioritizing FNA requirements according to what the government believes best meet its strategic objectives.
- Transition technology
- Collaborate and coordinate with other government programs, e.g., DARPA, DISA, OSD/NII to maximize the return on the government's objectives

Reordering the four investment priorities with respect to the long term view, places them in this order:

1. Make ready for the adoption of disruptive SI technology
2. Invest in the identification and proving of promising theories and approaches
3. Accelerate the maturity of emerging SI technologies that solve critical defense needs
4. Prove that an emerging or maturing commercial SI technologies can improve department functions, offer cost savings, and/or reduce risks

Integrating these relative priorities against filling FY07/08 gaps from the FNA would suggest the following choice of activity targets for potential longer term investment:

- Invest in SI infrastructure:
    - Ontological repository services for NCES
    - Ontological mediation services for NCES
    - Tools to assist COIs with ontology migration

- Invest in basic research toward approaches towards:
  - Quality of service
- Accelerate via contracted effort capabilities to develop:
  - Security and trust
  - Schema ontology maps
  - Ontology alignment
  - Knowledge design patterns
- Prepare to prove via demonstration activities:
  - All available SI capabilities

Considering the availability, types of funding available, and present FYDP funds alignment, it is now possible to develop an outyear investment and activity portfolio approach that appropriately mixes short and longterm objectives with salient demonstration opportunities (e.g., the OIM/JBI demo roadmap, JEFX08).

8.2.2 SI Non-Material Solution Roadmap

Non-material solutions must be considered in parallel with the development and adoption of solution technologies. This section discusses potential non-material solutions that must accompany or presage the material technology roadmap.

8.2.2.1 Doctrine and Policy

The FNA concluded that there is currently a gap in normative guidance to the department instructing that COIs and other programs should begin to assemble semantic metadata as well as structural metadata. It would help if such a policy was promulgated at around the time tools to assist the COIs (and hence a method) become available. Prior to the arrival of capable tools, there are still things that COIs can do to preserve the domain knowledge that will be used to form ontologies. Ideal policies to consider influencing are the next revisions to the Guidance for Implementing Net-Centric Data Sharing (presently DoD Instruction 8320.02G) and the emerging Information Sharing Environment from DNI (P32, P34).

As semantic solutions to information assurance become available, it is possible that instructions on information assurance policy will also have to be amended.

8.2.2.2 Organization

In the FNA we stated that no gaps presently exist of the organizational category. As SI is reduced to practice however, the USAF may want to consider an organizational solution to centrally managing its semantic infrastructure or distributing this responsibility out to information management organizations at the commands, numbered air forces, expeditionary wings, or Air Operations Centers.

8.2.2.3 Training

The most pressing need for training at present is for those who are attempting to compile domain knowledge and wish to form it into useful ontologies. Many COIs have knowledge management savvy individuals on hand to assist with this activity but just as many do not. It is unlikely that personnel who are not trained in knowledge management will build effective ontologies without considerable trial and error. A partial solution to this issue would be to collect, document, and disseminate relevant KM domain experience, best practices, model ontologies, lessons learned, and when available, tools. To some degree, training beyond the theoretical concept level will have to wait for the emergence of stable semantic tools (e.g., knowledge formation and discovery). Training will remain an issue, and likely an open gap as the component technologies rapidly emerge. User training will become an issue as SI becomes operationalized as it requires skillsets to manage, maintain, and adapt SI artifacts that are not presently in wide abundance in the force structure.

8.2.2.4 Leadership

The FNA suggested that there is not now a leadership gap with respect to semantic interoperability but that this may not remain the case if SI technology does not soon deliver compelling return on investment. For this reason, our short term investment plan for SI material technology focused on demonstrating its worth and potential as early as is possible. Clearly, the way to keep leadership engaged is to show SI's positive impact on classic air and ISR operations use cases and at large, service-wide exercises such as JEFX. SI will require continued leadership buy-in and support as the O&M for SI may require the allocation of new infrastructural resources and possibly affect training and organizational structure.

8.2.2.5 Personnel

It is unclear at this point whether knowledge engineering needs to be considered as an enlisted or officer AFSC or whether its concepts can be appended to the curricula of current enlisted and officer technical training schools. It is likely that SI and knowledge engineering skills could also be made part of the information operations and intelligence operations career field curricula. The shortage of trained knowledge engineers will chiefly be felt as SI is operationalized although this is also a skillset that can initially be outsourced to contractors.

8.2.2.6 Facilities

As stated in the FNA, facilities – in the form of networked resources – will be needed to register and control the department's persistent SI resources. These facilities will need to be available at multiple security levels.

9 Conclusion and Recommendations

Ten years from now, machine-to-machine semantic interoperability will be as ubiquitous and as effective as an interoperability technology as modern RDBMS technology is today for storing structured information. Machines will seek out information and services throughout the GIG and efficiently exchange critical information with full provision for releasability, quality of service, and information assurance. Humans will not be made to perform the tedious transformations that they must now endure for interactions to freely occur between disparate domains such as strike, logistics, and public affairs. The role of humans in SI will be in the preparation and maintenance of domain knowledge ontologies and the definition of the rules and business practices that will govern machine-based interactions.

Before this vision can be made a reality, however, a fairly steep technology gradient must be overcome as the bulk of the semantic technology family is relatively immature. Despite its apparent immaturity, large mainstream corporations including Microsoft, Yahoo, and IBM are now exploiting semantic capabilities to flexibly manage, exchange, and describe networked content. Indeed the promise of the Semantic Web has over the course of three years given rise to a vibrant industry that is now the primary exponent of the technology.

DoD's desire to implement a global information grid capable of supporting the manifold advantages of network-centric warfare have necessarily lead it to seek solutions made possible by semantic technologies. DoD network-centric policy demands that warfighting domains make their domain knowledge discoverable, accessible, and understandable so that they might participate in the GIG. Furthermore, the government has embarked on a broad initiative to collect and describe domain knowledge via the community of interest governance process and by investment in core GIG enterprise services. These are foundational and important steps toward semantic enablement.

To date however, these initiatives have been fully taxed accomplishing the daunting task of collecting, assembling, and managing structural metadata that describes the structure and syntax of the schemas of many communities of interest and have not yet begun to collect the richer semantic metadata needed to enable machine-to-machine (M2M) interoperability. The solution to this dilemma involves more than just the technology needed to capture and persist knowledge as ontologies, it also requires new policies that urge knowledge capture, training to know what to capture and how to use ontologies, and a tool base to support domain knowledge management.

Semantic technologies in their own right also have challenges that must be addressed such as differences in meaning between ontologies and differences in the semantic depth of exchanged information. SI is a disruptive technology in that it replaces long-standing ways that interoperability has been effected within the USAF and as such requires consideration of adjustments that must be made to the present technical and human infrastructure. It also needs a systematic approach to identify where investment can be made to operationalize semantics in the service of the GIG.

This study has attempted to capture the stated and implied requirements for and placed on semantic interoperability and has assembled these requirements into a functional architecture. The study then assessed where semantic technologies can now deliver functionality and where they are immature. The study also examined where existing DoD programs have begun to develop and apply the technologies that address the functional requirements. The study continued with an objective assessment of the maturity of elements of semantic technology and the identification of DOTMLPF gaps that exist in the application of SI technologies to M2M interoperability and policy. The remaining section of this Capabilities Based Assessment recommended how these gaps might be filled by material and non-material solutions and recommends an investment roadmap to help the department reach net-centricity quicker.

Since the government can be an effective exponent in the growth and adaptation of SI, we present a roadmap that recommends investment strategies. The roadmap recommends initial investment in the demonstration of key SI capabilities that are approaching operational maturity. This recommendation is made to ensure continued stakeholder buy-in as well as to raise awareness of the potential of SI technology. In concert with this initial demonstration, we recommend that investment be made in semantic tools to enable and equip the COIs compiling domain knowledge. We also advocate basic research to accelerate the development of specific technologies that are now immature yet represent critical USAF capability needs - specifically those which provide information assurance, trust, and quality of service. We cast the roadmap's investment recommendations in short and long term strategies and suggest the relative priorities for each strategy.

The commercial forces behind the Semantic Web and semantic interoperability guarantee its rapid maturity and lasting technology support. The USAF must prepare for semantic enablement by formalizing the capture of domain knowledge and establishing the semantic infrastructure that will in a short time allow the realization of effective machine-to-machine semantic interoperability and the projection of aerospace operations into the global information grid.

Appendix A Glossary of Key Terms

Accessible – Tenet of DoD Data Strategy. A data asset is accessible when a human, system, or application may retrieve the data within the asset. Data assets may be made accessible by using shared storage space or web services that expose the business or mission process that generates data in readily consumable forms. (DoD Dir, 8320.02)

Community of Interest (COI) – A collaborative group of users who must exchange information in pursuit of their shared goals, interests, missions, or business processes and who therefore must have shared vocabulary for the information they exchange. COIs are organizing constructs created to assist in implementing net-centric information sharing. Their members are responsible for making information visible, accessible, understandable, and promoting trust – all of which contribute to the data interoperability necessary for effective information sharing. (DoD Dir, 8320.02-G)

Context - In human-machine interaction, a context is a set of information that could be used to define and interpret a situation in which agents interact. In the context-aware applications community, context is composed of a set of information for characterizing the situation in which humans interact with applications and the immediate environment (Dey, 1998). In artificial intelligence, context is what does not intervene directly in problem solving but constrains it (Brézillon, 1999a). In short, a collection of relevant conditions and surroundings that make a situation unique and comprehensible.

Controlled Vocabulary - a carefully selected list of words and phrases, which are used to tag units of information (document or work) so that they may be more easily retrieved by a search. (Amy Warner, *A Taxonomy Primer*)

Data - A representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means. (DoD Dir, 8320.02) Data and information are typically equivalent terms, information often referring to data that are specific to a problem or use.

Data Asset - Any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a web site that returns data in response to specific queries (e.g., www.weather.com) would be a data asset. A human, system, or application may create a data asset. (DoD Dir, 8320.02)

Domain - Subsets of a Mission Areas and representing a common collection of related, or highly dependent, information capabilities and services. Managing these related information capabilities and services within domains improves coordination, collaboration, integration, and consistency of processes and interfaces for information sharing. (DoD Dir, 8320.02)

Domain Knowledge – Knowledge pertinent to a particular domain such as air operations, time-sensitive targeting, or finance.

Explicit Knowledge – Knowledge that typically includes some or all of context for the knowledge typically in the form of domain identification or attribution. For example a “Russian T-80 tank” is a more explicit term than “tank” as it grounds the term to a threat domain.

Expressivity – The degree of abstraction to which knowledge concepts are expressed. Very simple expressivity is provided by languages like XML tags which simply annotate that a certain object is associated with a concept. Deeper expressiveness results when concept relationships, behaviors, and rules are also recorded – as in an ontology.

Folksonomy – Popular means to establish and represent user-defined collaborative taxonomy usually by means of tagging.

Functor – Functors are objects that model operations that can be performed. In their simplest form they are somewhat like function pointers: they allow a client to call an unknown method with a standard interface.

Global Information Grid (GIG) **-** The globally connected, end-to-end set of information capabilities, associated processes, and personnel for collecting, processing, storing, disseminating, and managing information on demand to warfighters, policy makers, and support personnel. (DoD Dir, 8320.02)

Graph – In the Semantic Web, facts are represented as graphs composed to label nodes and relations, called Direct Acyclic Graphs.

Knowledge Representation – Means chosen to explicitly represent knowledge concepts. The Web Ontology Language (OWL) is an example of a means to represent knowledge.

Metadata - Information describing the characteristics of data; data or information about data; or descriptive information about an entity's data, data activities, systems, and holdings. For example, discovery metadata is a type of metadata that allows data assets to be found using enterprise search capabilities. (DoD Dir, 8320.02)

Network-Centric or **Net-Centric -** Relating to or representing the attributes of net-centricity. Net-centricity is a robust, globally interconnected network environment (including infrastructure, systems, processes, and people) in which data is shared timely and seamlessly among users, applications, and platforms. Net-centricity enables substantially improved military situational awareness and significantly shortened decision making cycles. Net-Centric capabilities enable network-centric operations and Net-Centric Warfare. (DoD Dir, 8320.02)

Network-Centric Warfare (NCW) **-** An information superiority-enabled concept of operations that generates increased combat power by networking sensors, decision makers, and shooters to achieve shared awareness, increased speed of command, higher tempo of operations, greater lethality, increased survivability, and a degree of self-synchronization. In essence, NCW translates information superiority into combat power by effectively linking knowledgeable entities in the battlespace. (DoD Dir, 8320.02)

Ontology – 1) A model that represents a group of concepts within a particular domain, usage rules, and the relationships between the concepts. Ontologies can be used to reason about objects within a domain. 2) An explicit specification of how to represent the objects and concepts that exist in some area of interest and of the relationships that pertain among them. (DoD Dir, 8320.02-G)

Ontology Alignment – The mapping of domain ontologies to an ontology that represents a shared understanding of underlying systems. This closely matches semantic interoperability across communities of interest.

Predicate – The portion of a phase that describes something about the object of the phrase.

Reasoning - Reasoning is the ability to make inferences, and automated reasoning is concerned with the building of computing systems that automate this process. (Stanford Encyclopedia of Philosophy).

Schema - A diagrammatic representation, an outline, or a model. In relation to data management, a schema can represent any generic model or structure that deals with the organization, format, structure, or relationship of data. Some examples of schemas are (1) a database table and relational structure, (2) a document type definition (DTD), (3) a data structure used to pass information between systems, and (4) an XML schema document (XSD) that represents a data structure and related information encoded as XML. Schemas typically do not contain information specific to a particular instance of data. (DoD Dir, 8320.02-G)

Schema-to-Ontology Mapping (S-O Mapping) – The process or artifact of mapping a local ontology of a domain to its underlying databases and services. The result is an ontology that is populated by instances retrieved from the underlying databases and services. A resulting populated ontology is called a Knowledgebase. This closely matches semantic interoperability within communities of interest.

Semantics - The meaning of phenomenology as it is represented in computer machines and it is often used in contrast with *syntax*. The meaning or relationship of meanings of terms and expressions within a system. Semantics deals with human interpretations according to their understanding of the world, and therefore, is context dependent.

Semantic Heterogeneity – Condition when more than one ontologies are considered. It is possible in this state to have terms in conflict.

Semantic Interoperability - The ability of information to flow between systems on the basis of shared, pre-established, and negotiated meanings of terms and expressions such that information is accurately and automatically interpreted by the receiving system.

Semantic Metadata – Information about a data asset that describes or identifies characteristics about that asset that convey meaning or context (e.g., descriptions, vocabularies, taxonomies). (DoD Dir, 8320.02-G)

Structural Metadata – Information provided about a data asset that describes the internal structure or representation of a data asset (e.g., database field names, schemas, web service tags). (DoD Dir, 8320.02)

Syntactic - Referring to the syntax or form of an object, not its meaning.

Tacit Knowledge – Knowledge that is understandable within a predefined context. Example, the term “fires” and “force” in the context of Army land combat operations can mean very different things when removed from that context.

Taxonomy – 1) Less broad than ontologies, taxonomies record parent-child relationships or the membership of elements in classes. 2) Provides categorizations of related terms. In doing so, they make use of “class/subclass” relationships (i.e., they are hierarchical in conveying the relationships between categories). Taxonomies are important to ensuring that searches of discovery metadata and content are targeted. (DoD Dir, 8320.02-G)

Triple – A subject/predicate/object relationship recorded to define a concept or a relationship between concepts. Example: “Cardinal/is a/Bird.” The Resource Descriptor Framework (RDF) for example, uses triples to express concepts. Ontologies can be built from facts and assertions encoded as triples.

Understandable - Tenet of DoD Data Strategy. Capable of being comprehended in terms of subject, specific content, relationships, sources, methods, quality, spatial and temporal dimensions, and other factors. (DoD Dir, 8320.02)

Visible - Tenet of DoD Data Strategy. Able to be seen, detected, or distinguished and to some extent characterized by humans and/or IT systems, applications, or other processes. (DoD Dir, 8320.02)

Vocabulary - Represents agreements on the terms and definitions common to a community of interest, including data dictionaries. For example, one community of interest might define the term “tank” to mean a pressurized vessel, whereas another might define “tank” to mean a tracked vehicle. Both definitions are acceptable, but the user must understand these definitions, and their context, to properly use the data. (DoD Dir, 8320.02-G)

Appendix B List of Consulted Documents

Policy and Doctrine Documents

Ref	Serial	Date	Title
P1	ASD (C3I) memorandum	20-Mar-97	Secret and Below Interoperability (SABI)
P2	AsstSecDef memorandum	21-May-02	Department of Defense Public Key Infrastructure (PKI)
P3	CJCSI 3170.01E	11-May-05	Joint Capabilities Integration and Development System
P4	CJCSI 6212.01D	8-Mar-06	Interoperability and Supportability of Information Technology and National Security Systems
P5	CJCSM 3500.04C	1-Jul-02	Universal Joint Task List
P6	DoD Directive 3222.3	20-Aug-90	Department of Defense Electromagnetic Compatibility Program
P7	DoD Directive 4630.5	11-Jan-02	Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)
P8	DoD Directive 4650.1	24-Jun-87	Management and Use of the Radio Frequency Spectrum
P9	DoD Directive 5000.1	12-May-03	The Defense Acquisition System
P10	DoD Directive 5100.35	10-Mar-98	Military Communications-Electronics Board (MCEB)
P11	DoD Directive 8100.1	19-Sep-02	Global Information Grid (GIG) Overarching Policy
P12	DoD Directive 8100.2	14-Apr-04	Use of Commercial Wireless Devices, Services, and Technologies in the Department of Defense (DoD) Global Information Grid (GIG)
P13	DoD Directive 8320.02-G	1-Apr-06	Guidance for Implementing Net-Centric Data Sharing
P14	DoD Directive 8320.2	2-Dec-04	Data Sharing in a Net-Centric Department of Defense
P15	DoD Directive 8500.1	24-Oct-02	Information Assurance
P16	DoD Instruction 4630.8	30-Jun-04	Procedures for Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)
P17	DoD Instruction 5200.40	30-Dec-97	DOD Information Technology Security Certification and Accreditation Process (DITSCAP)
P18	DoD Instruction 8500.2	6-Feb-03	Information Assurance (IA) Implementation
P19	DoD Memorandum	3-Apr-03	DoD Net-Centric Data Management Strategy: Metadata Registration (MID 905)
P20	DoD Memorandum	9-May-03	DoD Net-Centric Data Strategy
P21	DoD Regulation 5000.2-R	30-Oct-02	Interim Defense Acquisition Guidebook (formerly DOD Regulation 5000.2-R, 5 April 2002)
P22	DODAF v1.0	Oct-03	DoD Architectural Framework v1.0

P23	DODI 8110.1	6-Feb-04	Multinational Information Sharing Networks Implementation
P24	EO 13338	25-Oct-05	Further Strengthening the Sharing of Terrorism Information to Protect Americans
P25	GIG-A v2.0	13-Aug-02	Global Information Grid (GIG) Architecture Version 2.0
P26	IMSP v2.0	1-Oct-99	DoD Information Management (IM) Strategic Plan, Version 2.0
P27	JROCM 134-01	30-Aug-01	Global Information Grid Capstone Requirements Document
P28	JTA v6.0	24-Nov-03	DOD Joint Technical Architecture Version 6.0
P29	LISI	30-Mar-98	Levels of Information Systems Interoperability
P30	MCEB Pub 1	1-Mar-02	Organization, Mission and Functions Manual
P31	NSTISSP Number 11	5-Jun-03	National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products
P32	ODNI	28-Nov	Information Sharing Environment Implementation Plan
P33	Public Law No. 108-458	17-Dec-04	Intelligence Reform and Terrorism Prevention Act of 2004
P34	Requirement 1	Dec-06	Leveraging Ongoing Information Sharing Efforts in the Development of the ISE Substantial Progress
P35		9-May-03	Department of Defense, Net-Centric Data Strategy
P36		19-Jul-06	ESC Implementation Roadmap for Net-Centric Data Strategy
P37		15-Nov-05	Implementing the NCDS using COIs
P38		16-Dec-04	JMBC2 Data Strategy
P39	OASD Memorandum	6-Jan-04	NII CIO Network Centric Attributes List
P40	OASD Memorandum	1-Dec-05	NII CIO Network Centric Checklist
P41	OASD	4-Nov-04	Net-Centric Operations and Warfare (NCOW) Reference Model
P42	OASD Memorandum	24-Oct-03	Net-Centric Data Strategy: Visibility - Tagging and Advertising Data Assets with Discovery Metadata
P43	JROCM 199-04	29-Oct-04	CJCS Memorandum CM-2040-04 Assignment of Warfighting Mission Area Responsibilities to Support GIG ES

Community of Interest Materials

Ref	Date	Title
C1	24-Oct-06	Air Operations COI Overview - Common Mission Definition
C2	2-Oct-06	Blue Force Tracking COI Intro Brief
C3	16-Dec-04	Blue Force Tracking Data Strategy
C4	26-Jul-06	COI 101 - Enabling Information Sharing and Agility through Communities of Interest
C5	26-Jul-06	COI 201 - COI Basics

C6	26-Jul-06	COI 301 - COI Pilot Process Overview
C7	26-Jul-06	COI 401 - Approach for Defining and Validating COI Vocabulary
C8	1-Oct-05	Joint Track Management COI Analysis and Roadmap
C9	1-Oct-06	Joint Track Management Program Overview
C10	12-Oct-06	Maritime Domain Awareness COI Brief
C11	1-Nov-06	Strike COI Best of Breed Approach
C12	6-Oct-06	Strike COI DWG Final Minutes
C13	5-Oct-06	Strike COI Final Data Management Working Group Status Brief
C14	24-Oct-06	Strike COI JC3IEDM Brief
C15	26-May06	TST COI Data Management Panel Product Introduction

Data Models, Schema, and Vocabularies

Ref	Date	Title
D1	19-Aug-06	Blue Force Tracking COI Information Exchange Standard
D2	28-Sep-05	Common Battlespace Object Data Model v2.0 (XSD)
D3	28-Sep-05	Common Battlespace Object Design Description
D4	25-Oct-06	Common Battlespace Object revised data definitions
D5	3/27/2006	Common Mission Definition 1.0 Schema (XSD, XML, WSDL)
D6	29-Jul-05	DDMS - DoD Discovery Metadata Standard
D7	27-Jul-05	DDMS v1.3 Schema Set (XML, XSD, etc)
D8	27-Jul-05	DDMS v1.3 User's Guide
D9	20-Sep-06	JC3IEDM Domain Values
D10	10-May06	JC3IEDM v3.0 Model
D11	1-Sep-06	Joint Track Management Data Model
D12	1-Sep-06	Joint Track Management Logical Data Model
D13	10-Jul-06	Joint Track Management Logical Data Model Description
D14	1-Oct-06	Joint Track Management Status Brief
D15	18-Aug-06	Maritime Domain Awareness Schema Set (XML, XSD, etc.)
D16	Jul-06	Operational Joint Architecture Working Group Core Data Model
D17	29-Aug-06	Strike COI Model (XML, XSD, etc)
D18	20-Jul-06	Strike COI Spiral I Schema List
D19	4-Aug-06	Strike COI Taxonomy Draft
D20	15-Oct-06	Strike COI UML Vocabulary and Use Case
D21	24-Oct-06	Strike COI Updated UML
D22	1-Aug-06	TST COI Auxiliary Schema (XML)
D23	1-Aug-06	TST COI Base Schema (XML)
D24	1-Aug-06	TST COI Common Schema (XML)
D25	1-Aug-06	TST COI Common Schema (XSD)
D26	1-Jul-06	TST COI TST Vocabulary

Misc Program Documentation

Ref	Date	Title
M1	12/21/2005	AFSAB Domain Integration Study
M2	28-Sep-05	Common Battlespace Object Executive White Paper
M3	19-Oct-05	Common Battlespace Object Lessons Learned
M4	1-Jan-05	Cursor on Target - Situation Awareness Using
M5	25-Apr-05	Cursor on Target Developer's Guide v4
M6	23-Jun-04	Cursor on Target Program Brief
M7	12-Nov-03	GIG Enterprise Services
M8	1-Oct-05	Joint Track Management Study

M9	27-Oct-06	Minnowbrook QED WG Findings
M10	27-Oct-06	Minnowbrook SI WG Findings
M11	27-Oct-06	Minnowbrook Tactical Infospace Dominance WG Findings
M12	24-May-06	NCES Capability Development Document
M13	25-Apr-05	NCES Channel Administration Portlet Users Guide v0.4.3
M14	31-Aug-05	NCES Content Discovery Datasource Administration Portlet Users Guide
M15	31-Aug-05	NCES Content Discovery Federated Search Portlet Users Guide v0.4.3
M16	25-Apr-05	NCES Mediation Core Enterprise Services SDK v0.5.0
M17	25-Apr-05	NCES Messaging Store Portlet Users Guide v0.4.3
M18	16-May-05	NCES Security Service SRS v0.6.0
M19	15-Nov-06	NCES Program Overview
M20	30-Sep-05	NCES Service Data Gathering Checklist v1.3
M21	2-Oct-04	NCES Service Discovery Core Enterprise Services CONOPs v0.4
M22	29-Jun-05	NCES Service Discovery Publishing Services to a UDDI Registry
M23	23-May-05	NCES Service Security Design Specification
M24	23-May-05	NCES Service Security Interface Specification
M25	21-Nov-06	NCES Software Baseline
M26	11-Oct-06	Net-Centric Enterprise Services Net-Enabled Command Capability
M27	1-Mar-04	Situational Awareness Using Cursor on Target

Misc Use Case Documents

Ref	Date	Title
U1	25-Oct-05	Common Battlespace Object VMF Use Case
U2	9-Aug-06	SPAWAR TST Thread Model
U3	31-Jul-06	Strike COI Use Case v4
U4	1-Jun-05	Time Sensitive Targeting Mission Thread v0.5

Appendix C Requirements Taxonomy

This taxonomy is structured primarily by the central tenets of the DoD Net-Centric Data Strategy (DoD Instruction 8320.02). The categories derive from an investigation of policy and doctrine information as well as the needs of ongoing programs and governance and finally the implications of adopting a semantic solution to the interoperability problem. To the right of each is the functional architecture category that each is assigned to. These include:

B&A -	Binding and Access
DK-	Domain Knowledge
Mig -	Migration
Pers -	Persistence
P&D -	Publish and Discover
QoS -	Quality of Service
Sec -	Security
Tool -	Tools
W&P -	Workflow and Planning

Visibility and Awareness (Is an information resource discoverable?)

- Discovery / Publishing / Binding
  - Services
    - Find — P&D
    - Bind — B&A
    - Publish — P&D
    - Publish Service Description — P&D
    - Identify utilization of DDMS — P&D
    - Indicate whether service is providing access to source or a copy — P&D
    - Indicate if in storage vs service accessible — P&D
    - Indicate NCOW protocol standard used — P&D
    - Indicate public / private intent — P&D
    - Identify domain — P&D
    - Indicate minimum anticipated support for edge devices — P&D
    - Identify IPv4 or IPv6 — P&D
    - Provide registry services — P&D
    - Provide catalog services — P&D
    - Register with provided service — P&D
    - Method to bind requestor to service — B&A
    - Production of a data dictionary — P&D
  - Schema
    - Find — P&D
    - Bind — B&A
    - Publish — P&D

- Instance
  - Publish Exchangeable Content — P&D
  - Discover Exchangeable Content — P&D
- Mappings and correspondences
  - Find — P&D
  - Bind — B&A
  - Publish — P&D
- Ontology
  - Find — P&D
  - Bind — B&A
  - Publish — P&D
- Nodes (e.g., ISR)
  - Find — P&D
  - Bind — B&A
  - Publish — P&D

- Tagging — DK

Accessibility (Is it available to me on the network and do I have the tools to use it?)

- Throughput — QoS
- Perform Low Bandwidth Exchange — QoS
- Manage discontinuous operations — QoS
- Security — Sec
  - extend security context — Sec
  - mediate security assertions — Sec
  - method of authentication — Sec
  - accommodate PKI — Sec
  - accommodate XML signature — Sec
  - accommodate XML encryption — Sec
  - insert security assertions — Sec
  - Mark for security / releasability — Sec
  - Mark for handling caveats — Sec
  - Support for security guards — Sec
  - Adjudicate access rights of user from service — Sec
- Persistence — Pers
  - Store semantic content — Pers
- Workflow and Planning — W&P
  - Scheduling service invocations — W&P
  - Chaining together services (planning) — W&P
  - Prioritization of information delivery — W&P

Understandability (Can I intelligibly use it? Do I understand the semantics?)

- Geospatial/temporal — DK
  - Accommodate positional information — DK
  - Accommodate temporal info — DK
- Controlled/uncontrolled vocabularies — DK

- Knowing role of information (command, inform, discover, describe, acknowledge) DK
- Knowing role of sender and receiver DK
- Handling uncertainty of facts DK
- Managing Ambiguity DK
- Resource identity within domain (e.g., FSN, BE Number) DK
- Object classification criteria/methodology DK
- Managing Complexity DK
- Capturing user intent DK
- Transforming information according to context Mig
- Identifying contradictions (e.g., mission [target] conflicts with policy [no-hit list]) QoS

Interoperability (Can it be combined or compared with other information?)

- Mediation Mig
    - Map concept between ontologies Mig
- Semantic Similarity Mig
    - Compare ontologies Mig
- Multi-source data fusion Mig
- Reasoning Mig

Jointness (Are the semantics oriented to the joint warfighter or a specific group?)

- Internationalization DK

Quality (Can the data be trusted to be accurate and reliable?)

- Quality of Service QoS
    - Checksum QoS
    - Acknowledgement QoS
    - Verify content QoS
    - Support to a Service Level Agreement QoS
- Versioning and Reuse Mig
    - Mark version Mig
    - Change impact amelioration QoS
- Unique identification Mig
    - Uniquely identify schema Mig
- Pedigree QoS
- Provenance QoS
- Trust QoS
- Validation of ontology QoS
- Reliability/Availability of resources QoS

Management Tools

- Methodology of ontology development Tool
    - Patterns Tool
- Infosphere creation and self identification Tool

• Structuring and controlling infosphere	Tool
• Establishing initial info sets	Tool
• Dissolution	Tool

Other Policy-Based Requirements

• If deals with terrorism info must comply with IRTPA	Other
• Support enterprise search	P&D
• Shall use DDMS and ISM	Other
• Compliance with the Net-Centric Operations and Warfare (NCOW) Reference Model (RM)	Other
• Compliance with Applicable Global Information Grid (GIG) Key Interface Profiles (KIP)	Other
• Compliance with DOD information assurance requirements	Sec
• Compliance with supporting integrated architecture products required to assess information exchange and use for a given capability	Other
• Information assurance via defense in depth approach	Sec
• Compliance with applicable GIG Key Interface Profiles (KIPs)	Other
• Comply with the most current version of the DOD DISR	Other
• Compliance with DOD Directive 8500.1 and DOD Instruction 8500.2, and with Phase 1 Definition of the DITSCAP (DOD Instruction 8500.40)	Sec
• Support JITC testing	QoS

Appendix D FNA: SI Technologies

Technology Name	Description
3store	A core C library that uses MySQL to store its raw RDF data and caches, forming an important part of the infrastructure required to support a range of knowledgeable services
4Suite 4RDF	The 4Suite 4RDF an open-source platform for XML and RDF processing implemented in Python with C extensions
ActiveRDF	ActiveRDF is a library for accessing RDF data from Ruby programs. It can be used as data layer in Ruby-on-Rails. You can address RDF resources, classes, properties, etc. programmatically, without queries
Adaptiva	A user-centered ontology building environment, based on using multiple strategies to construct an ontology, minimizing user input by using adaptive information extraction
Aduna Metadata Server	The Aduna Metadata Server automatically extracts metadata from information sources, like a file server, an intranet or public web sites. The Aduna Metadata Server is a powerful and scalable store for metadata
AJAX Client for SPARQL	AJAX Client for SPARQL is a simple AJAX client that can be used for running SELECT queries against a service and then integrating them with client-side Javascript code
AKT-Bus	An open, lightweight, Web standards-based communication infrastructure to support interoperability among knowledge services.
AllegroGraph	Franz Inc's AllegroGraph is a system to load, store and query RDF data. It includes a SPARQL interface and RDFS reasoning. It has a Java and a Prolog interface
Almo	An ontology-based workflow engine in Java
Altova SemanticWorks	Visual RDF and OWL editor that auto-generates RDF/XML or nTriples based on visual ontology design
Amilcare	An adaptive information extraction tool designed to support document annotation for the Semantic Web.
ANNIE - Open Source Information Extraction	An open-source robust information extraction system
Aperture	Aperture is a Java framework for extracting and querying full-text content and metadata from various information systems (e.g. file systems, web sites, mail boxes) and the file formats (e.g. documents, images) occurring in these systems
Applications of FCA in AKT	Formal Concept Analysis (FCA) is used in a variety of application scenarios in AKT in order to perform concept-based domain analysis and automatically deduce a taxonomy lattice of that domain.
ARC	ARC is a lightweight, SPARQL-enabled RDF system for mainstream Web projects. It is written in PHP and has been optimized for shared Web environments
Armadillo	Exploits the redundancies apparent in the Internet, combining many information sources to perform document annotation with minimal human intervention.
ATLAS	ATLAS (Architecture and Tools for Linguistic Analysis Systems) is a joint initiative of NIST, MITRE and the LDC to build a general purpose annotation architecture and a data interchange format. The starting point is the annotation

	graph model, with some significant generalizations
AutoSemantix	AutoSemantix is a round-trip code generation tool designed to streamline the creation of Semantic Web applications for the Java platform
BBN OWL Validator	BBN OWL Validator
Beagle++	Beagle++ is an extension to the Beagle search tool for the personal information space. Beagle++ now makes that search semantic, moving towards a vision of the Semantic Desktop
Bossam	Bossam, a rule-based OWL reasoner (free, well-documented, closed-source)
Brahms	Brahms is a fast main-memory RDF/S storage, capable of storing, accessing and querying large ontologies. It is implemented as a set of C++ classes
BrownSauce	The BrownSauce RDF browser is a project to aggregate and present arbitrary RDF data in as pleasing a manner as possible, that is a 'semantic web browser'. Brownsauce is a local http server; however it should be trivial to add other front-ends
CARA	CARA (*CA*RMEN *R*DF *A*PI) provides an API for the Resource Description Framework (RDF). The API is based on the graph model of RDF, supports in-memory and persistent storage and includes an RDF Parser
CASheW-s Engine	The purpose of this project is to facilitate the composition of semantic web services. It consists of two parts, of which this is one
COCKATOO	A knowledge acquisition tool which can be used to produce a set of cases for use with a Case-Based Reasoning system.
Compendium	Compendium is a semantic, visual hypertext tool for supporting collaborative domain modeling and real time meeting capture
ConRef	A service discovery system which uses ontology mapping techniques to support different user vocabularies
ConcepTool	A system to model, analyze, verify, validate, share, combine, and reuse domain knowledge bases and ontologies, reasoning about their implication.
Corese	Corese stands for Conceptual Resource Search Engine. It is an RDF engine based on Conceptual Graphs (CG) and written in Java. It enables the processing of RDF Schema and RDF statements within the CG formalism, provides a rule engine and a query engine accepting the SPARQL syntax
cwm	The Closed World Machine (CWM) data manipulator, rules processor and query system mostly using the Notation 3 textual RDF syntax. It also has an incomplete OWL Full and a SPARQL access. It is written in Python
D2R MAP Processor	D2R MAP is a declarative language to describe mappings between relational database schemata and OWL ontologies. This D2R processor implements the D2R mapping language and exports data from a relational database as RDF, N3, N-TRIPLES or as Jena models
D2R Server	D2R Server, turns relational databases into SPARQL endpoints, based on Jena's Joseki
DBIN	DBin brings the Semantic Web to the end users. By joining P2P groups and communities, users can annotate any topic or subject of interest and enjoy browsing and editing in a semantically rich environment.
DOSE	A distributed platform for semantic annotation
Drive	Drive is an RDF parser written in C# for the .NET platform
ekoss.org	A collaborative knowledge sharing environment where model developers can submit advertisements
Euler	Euler is an inference engine supporting logic based proofs. It is a backward-chaining reasoner enhanced with Euler path detection. It has implementations in Java, C#, Python, Javascript and Prolog. Via N3 it is interoperable with W3C Cwm

Exteca	The Exteca platform is an ontology-based technology written in Java for high-quality knowledge management and document categorization. It can be used in conjunction with search engines
ExtrAKT	ExtrAKT is a tool for extracting ontologies from Prolog knowledge bases.
F-Life	F-Life is a tool for analyzing and maintaining life-cycle patterns in ontology development.
FaCT++	FaCT++ is an OWL DL Reasoner implemented in C++
Floodsim	A prototype system which demonstrates the benefits of applying semantically rich service descriptions (expressed using Semantic Web technologies) to Web Services.
FOAF-o-matic	Online Friend OF A Friend generator
FOAM	Framework for ontology alignment and mapping
Foxtrot	Foxtrot is a recommender system which represents user profiles in ontological terms, allowing inference, bootstrapping and profile visualization
Fresh Framework	Fresh Framework is a CMS designed for the Semantic Web, with WYSIWYG page editing, RDF summaries of profiles and news, and countless other quality features you expect to find in a CMS
GNOWSYS	GNOWSYS, Gnowledge Networking and Organizing System, is a web based hybrid knowledge base with a kernel for semantic computing. It is developed in Python and works as an installed product in ZOPE
Graphl	Graphl is a generic graph visualization and manipulation tool written in Java.
Groove	Graph transformation, model transformation, object-oriented verification, behavioral semantics
GrOWL	Open source graphical ontology browser and editor
HAWK	OWL repository framework and toolkit
Haystack	Haystack is a tool designed to let individuals manage all their information in ways that make the most sense to them.
HELENOS	A Knowledge discovery workbench for the semantic Web
hMAFRA (Harmonize Mapping Framework)	hMAFRA is a set of tools supporting semantic mapping definition and data reconciliation between ontologies. The targeted formats are XSD, RDFS and KAON
I-X Process Panels	The I-X tool suite supports principled collaborations of human and computer agents in the creation or modification of some product
IBM Semantics Toolkit	BM Semantics Toolkit is designed for storage, manipulation, query, and inference of ontologies and corresponding instances. A major purpose is to establish an end-to-end ontology engineering environment tightly integrated with dominant Meta- Object Facility (MOF)-based modeling and application development tools. The semantics toolkit contains three main components (Orient, EODM, and RStar), which are designed for users of different levels.
Identify Knowledge Base	Identify-Knowledge-Base is a tool of Topic Identification about Knowledge Base
IF-Map	IF-Map is an Information Flow based ontology mapping method. It is based on the theoretical grounds of logic of distributed systems and provides an automated streamlined process for generating mappings between ontologies of the same domain.
IkeWiki	IkeWiki is a new kind of Wiki (a so-called Semantic Wiki") developed by Salzburg Research
Internet Reasoning Service	The Internet Reasoning Service provides a a number of tools which supports the publication, location, composition and execution of heterogeneous web services, specified using semantic web technology
IODT	IBM's toolkit for ontology-driven development

IsaViz	IsaViz is a visual authoring tool for browsing and authoring RDF models represented as graphs. Developed by Emmanuel Pietriga of W3C and Xerox Research Centre Europe
Jambalaya	Protégé plug-in for visualizing ontologies
Jastor	Open source Java code generator that emits Java Beans from ontologies
Javascript RDF/Turtle parser	Javascript RDF/Turtle parser, can be used with Jibbering
Jena	Jena is a Java framework to construct Semantic Web Applications. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL and includes a rule-based inference engine. It also has the ability to be used as an RDF database via its Joseki layer.
JessTab	JessTab is a plug-in for Protégé that allows you to use Jess and Protégé together. JessTab provides a Jess console window where you can interact with Jess while running Protégé.
Jibbering	Jibbering, a simple javascript RDF Parser and query thingy
Joseki	Jena's Joseki layer offers an RDF Triple Store facility with SPARQL interface
JRDF	JRDF Java RDF Binding is an attempt to create a standard set of APIs and base implementations to RDF using Java. Includes a SPARQL GUI.
KAON	Open source ontology management infrastructure
KAON2	KAON2 is an an infrastructure for managing OWL-DL, SWRL, and F-Logic ontologies. it is capable of manipulating OWL-DL ontologies; queries can be formulated using SPARQL
Kazuki	Generates a java API for working with OWL instance data directly from a set of OWL ontologies
KIM Platform	KIM is a software platform for the semantic annotation of text, automatic ontology population, indexing and retrieval, and information extraction from Ontotext
KnoZilla	
Knowledge Broker	The knowledge broker addresses the problem of knowledge service location in distributed environments.
knowledgeSmarts	knowledgeSmarts is a Java framework to construct Semantic Web Applications. It supports geospatial and temporal reasoning and allows real-time integration of a wide range of database. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL, and OWL-S. It has a pluggable architecture to rule-based inference engines and DL reasoners.
Kowari	Open source database for RDF and OWL
KRAFT - I-X TIE	Supports collaboration among members of a virtual organization by integrating workflow and communication technology with constraint solving.
Kraken	Kraken is an application for managing knowledge objects, which can be documents, remote or locally cached Web pages, personal information, to do list items, appointments, and so on. It is especially useful for researchers or students to manage their information.
LinKFactory	Language & Computing's LinKFactory is an ontology management tool, it provides an effective and user-friendly way to create, maintain and extend extensive multilingual terminology systems and ontologies (English, Spanish, French, etc.).
Longwell	Longwell is a web-based RDF-powered highly-configurable faceted browser
Lucene	Apache Lucene is a high-performance, full-featured text search engine library written entirely in Java. It is a technology suitable for nearly any application that requires full-text search, especially cross-platform. It is open source

LuMriX	A commercial search engine using semantic Web technologies
Machinese Syntax	Machinese Syntax provides a full analysis of texts by showing how words and concepts relate to each other in sentences - still with very competitive speed and accuracy. Machinese Syntax helps analytic applications understand text beyond the level of words, phrases and entities: also their interrelations (such as events, actions, states and circumstances); from Connexor
MAFRA Toolkit	Ontology MApping FRAmework Toolkit allows to create semantic relations between two (source and target) ontologies, and apply such relations in translating source ontology instances into target ontology instances
Magpie	Magpie supports the *interpretation* of web documents through on-the-fly ontologically based enrichment. Semantic services can be invoked either by the user or be automatically triggered by patterns of browsing activity
MatrixBrowser Visualization Kit	The MatrixBrowser project presents a new approach for visualizing and exploring large networked information structures which may represent, for instance, linked information resources or metadata structures such as ontologies
Melita	Melita is a semi-automatic annotation tool using an Adaptive Information Extraction engine (Amilcare)to support the user in document annotation
MetaDesk	MetaDesk is an RDF authoring tool that emphasizes entry of facts, rather than construction of ontologies. MetaDesk places no restrictions on vocabulary- users can invent terms on-the-fly, which the system converts into underlying RDF structures.
MetaMatrix	Semantic vocabulary mediation and other tools
Metatomix	Commercial semantic toolkits and editors
MindRaider	Open source semantic Web outline editor
MnM	MnM is an annotation tool which provides both automated and semi-automated support for annotating web pages with semantic contents. MnM integrates a web browser with an ontology editor and provides open APIs to link to ontology servers and for integrating information extraction tool
Model Futures OWL Editor	Simple OWL tools, featuring UML (XMI), ErWin, thesaurus and imports
Morla	Editor of RDF documents that allows you to manage more RDF documents simultaneously, visualize graphs, and use templates for quick writing. You can import RDFS documents and use their content to write new RDF triples. Templates are also RDF documents, and they make Morla easily personalizable and expandable. You can also use Morla as an RDF navigator, browsing the RDF documents present on the Internet exactly as you are used to doing with normal browsers
Mulgara	The Mulgara Semantic Store is an Open Source, massively scalable, transaction-safe, purpose-built database for the storage and retrieval of RDF, written in Java. It is an active fork of Kowari
Muskrat-II	Given a set of knowledge bases and problems solvers, the Muskrat system will try to identify which knowledge bases could be combined with which problems solvers to solve a given problem.
MyPlanet	MyPlanet allows users to create a personalized version of a web based newsletter using an ontologically based profile.
Net OWL	Entity extraction engine from SRA International
NMARKUP	NMARKUP helps the user build ontologies by detecting nouns in texts and by providing support for the creation of an ontology based on the entities extracted.
Nokia Semantic Web	An RDF based knowledge portal for publishing both authoritative and third

Server	party descriptions of URI denoted resources
Nuin BDI Agent Engine	A Java BDI agent engine for semantic web agents
Oink	OINK is a browser for RDF data. OINK queries data in an RDF triple store, and renders it as XHTML pages (essentially, one page per each node in the graph, on demand). This allows one to view RDF (and OWL) data in a very clear, intuitive way. OINK is built on top of Wilbur
OMCSNet-WordNet	The OMCSNet-WordNet project aims to improve the quality of the OMCSNet dataset by using automated processes to map WordNet synonym sets to OMCSNet concepts and import additional semantic linkage data from WordNet. It is based on OMCSNet 1.2, a semantic network and inference toolkit written in Python/Java. OMCSNet currently contains over 280,000 separate pieces of common sense information extracted from the raw OMCS dataset. This project is also based on WordNet, an online lexical reference system that in recent years has become a popular tool for AI researchers
ONTOCOPI	A tool which uncovers Communities Of Practice by analyzing the connectivity of instances in the 3store knowledge base.
OntoEdit/OntoStudio	Engineering environment for ontologies
Ontology Organizer	A DAML+OIL ontology editor with constraint propagation functionality to ensure that constraints applied to properties and restrictions are correctly propagated through an ontology, and datatype management functionality for manipulating custom datatypes
OntoMat Annotizer	Interactive Web page OWL and semantic annotator tool
OntoPortal	Enables the authoring and navigation of large semantically-powered portals
OpenLink Data Spaces (ODS)	ODS is a distributed collaborative application platform for creating Semantic Web applications such as: blogs, wikis, feed aggregators, etc., with built-in SPARQL support and incorporation of shared ontologies such as SIOC, FOAF, and Atom OWL. ODS is an application of OpenLink Virtuoso and is available in Open Source and Commercial Editions.
Oracle Spatial 10g	Oracle Spatial 10g includes an open, scalable, secure and reliable RDF management platform
Oyster	Peer-to-peer system for storing and sharing ontology metadata
OWL API	A Java interface and implementation for the W3C Web Ontology Language (OWL), used to represent Semantic Web ontologies. The API is focused towards OWL Lite and OWL DL and offers an interface to inference engines and validation functionality
OWL Consistency checker	OWL Consistency checker (based on Pellet)
OWL-DL Validator	WonderWeb OWL-DL Validator
OWLJessKB	OWLJessKB is a description logic reasoner for OWL. The semantics of the language is implemented using Jess, the Java Expert System Shell. Currently most of the common features of OWL lite, plus some and minus some
OWLLib	This is PHP library for accessing OWL files. OWL is w3.org standard for storing semantic information
OWLIM	OWLIM is a high-performance semantic repository, packaged as a Storage and Inference Layer (SAIL) for the Sesame RDF database
OWLViz	OWLViz is visual editor for OWL and is available as a Protégé plug-in
Pedro	Pedro is an application that creates data entry forms based on a data model written in a particular style of XML Schema. Users can enter data through the forms to create data files that conform to the schema. They can use controlled vocabularies to mark-up text fields and have the application perform basic

	validation on field data
Pellet	Pellet is an open-source Java based OWL DL reasoner. It can be used in conjunction with both Jena and OWL API libraries; it can also be downloaded and be included in other applications
Piggy Bank	A Firefox-based semantic Web browser
Pike	A dynamic programming (scripting) language similar to Java and C for the semantic Web
Platypus Wiki	Platypus Wiki is an enhanced Wiki Wiki Web with ideas taken from Semantic Web. It offers a simple user interface to create a Wiki Page plus metadata according with W3C standards. It uses RDF/RDFS and OWL to create ontologies and manage metadata
POR	Protege+OWL+Ruby (POR) Utilities provides an ontology, a set of ruby classes and methods to simplify the development of Protege+OWL Ontology Driven applications. At the moment project is limited to JRuby
pOWL	Semantic Web development platform
Protégé	Open source visual ontology editor written in Java with many plug-in tools
Pytypus Wiki	Pytypus is a Semantic Web project. In Pytypus, RDF is the base of communication between agents in the semantic net. Every URI in the semantic net has its owner that rule its behavior
RACER	A collection of Projects and Tools to be used with the semantic reasoning engine RacerPro
RacerPro	RacerPro is an OWL reasoner and inference server for the Semantic Web
Raptor	The Raptor RDF parser toolkit is a free software / Open Source C library that provides a set of parsers and serializers that generate Resource Description Framework (RDF) triples by parsing syntaxes or serialize the triples into >a syntax. The supported parsing syntaxes are RDF/XML, N-Triples, Turtle, RSS tag soup including Atom 1.0 and 0.3, GRDDL for XHTML and XML. The serializing syntaxes are RDF/XML (regular, and abbreviated), N-Triples, RSS 1.0, Atom 1.0 and Adobe XM
Rasqual	Rasqal is a C library for querying RDF, supporting the RDQL and SPARQL languages. It provides APIs for creating a query and parsing query syntax. It features pluggable triple-store source and matching interfaces, an engine for executing the queries and an API for manipulating results as bindings. It uses the Raptor RDF parser to return triples from RDF content and can alternatively work with the Redland RDF library's persistent triple stores. It is portable across many POSIX systems
rdfabout.com's Validator	RDF/XML and N3 validator
RDF Filter	This program acts as a filter layer between SAX (The Simple API for XML) and the higher-level RDF (Resource Description Format), an XML-based object-serialization and metadata format. The RDF filter library is used by several RDF-based projects
RDF Gateway	Intellidimension's RDF Gateway is an RDF Triple database with RDFS reasoning and SPARQL interface
RDF InferEd	Intellidimension's RDF InferEd is an authoring environment with the ability to navigate and edit RDF documents
RDFizers	RDFIzers arew little conversion tools for converting a source file in a given format to RDF. RDFizers are provided for JPEG, MARC/MODS, OAI-PMH, OCW, EMail, BibTEX, Flat, Weather, Java, Javadoc, Jira, Subversion and Random. In addition, the project page has links to other third-party RDF converters for iCal, Palm, Outlook, RFC822, Garmin, EXIF, Fink, D2RQ,

	D2RMAP, XLS, CSV, XSD, XML and MPEG-7/CS
RDFLib	RDFLib, an RDF libary for Python, including a SPARQL API. The library also contains both in-memory and persistent Graph backends
RDFReactor	Access RDF from Java using inferencing
RDF Server	The RDF server of the PHP RAP environment
RDFStore	RDFStore is an RDF storage with Perl and C API-s and SPARQL facilities
RDFSuite	The ICS-FORTH RDFSuite open source, high-level scalable tools for the Semantic Web. This suite includes Validating RDF Parser (VRP), a RDF Schema Specific DataBase (RSSDB) and supporting RDF Query Language (RQL)
RDFX	RDFX is a suite of plug-ins for the Eclipse platform designed to encourage and facilitate experimentation of semantically enhanced applications
Redfoot	Redfoot is a hypercoding system which is being used to create a webized operating system and is also being used to create applications. It is built around the notion of an RDF Graph for persistence rather than a File Tree
Redland	The Redland RDF Application Framework is a set of free software libraries that provide support for RDF. It provides parser for RDF/XML, Turtle, N-triples, Atom, RSS; has a SPARQL and GRDDL implementation, and has language interfaces to C#, Python, Obj-C, Perl, PHP, Ruby, Java and Tcl
RelationalOWL	Automatically extracts the semantics of virtually any relational database and transforms this information automatically into RDF/OW
ReTAX+	ReTAX is an aide to help a taxonomist create a consistent taxonomy and in particular provides suggestions as to where a new entity could be placed in the taxonomy whilst retaining the integrity of the revised taxonomy (c.f., problems in ontology modeling).
Refiner++	REFINER++ is a system which allows domain experts to create and maintain their own Knowledge Bases, and to receive suggestions as to how to remove inconsistencies, if they exist.
Rhizome Wiki	Rhizome is a Wiki-like content management and delivery system that exposes the entire site including content, structure, and metadata as editable RDF. This means that instead of creating a site with URLs that correspond to a page of HTML, you can create URLs that represent just about anything. It was designed to enable non-technical users to create these representations in an easy, ad-hoc manner. For developers, this allows both content and structure to be easily repurposed and complex Web applications to be rapidly developed
Rx4RDF	Rx4RDF shields developers from the complexity of RDF by enabling you to use familiar XML technologies like XPath, XSLT and XUpdate to query, transform and manipulate RDF. Also included is Rhizome, a wiki-like application for viewing and editing RDF models
Seamark Navigator	Siderean's Seamark Navigator provides a platform to combine Web search pages with product catalog databases, document servers, and other digital information from both inside and outside the enterprise
Searchy	Searchy is a metasearch engine that is able to integrate information from a wide range of sources performing a semantic translation into RDF. It has a distributed nature and is specially suitable to integrate information across different organizations with a minimum coupling
SECO	SECO provides mediation services for Semantic Web data, comprising data acquisition and data integration mediators. A SECO mediator comprises an HTTP server, an RDQL parser, and means to fetch data via RDQL/HTTP. User interface and scutter can accept commands via HTTP GET, where the user interface serves HTML pages, and the scutter fetches a page

Semantic Annotation with MnM	MnM is a semantic annotation tool which provides manual, automated and semi-automated support for annotating web pages with 'semantics', i.e., machine interpretable descriptions.
Semantical	Open source semantic Web search engine
Semantic Engine	The Semantic Engine is a standalone indexer/search application. Mac OSX only; Windows and Linux versions are on their way
Semantic Explorer	The Semantic Explorer allows you to enter a search query and watch as the resulting sub-graph is laid out on screen, visually clustering documents and terms together. Mac OS X only
Semantic Tools for Web Services	Semantic Tools for Web Services is a set of Eclipse plug-ins that allow developers to insert semantic annotations into a WSDL document to describe the semantics of the input, output, preconditions, and effects of service operations. A second plug-in matches the description of the service or composition of services to that for which a developer is searching. This technology is part of the Emerging Technologies Toolkit (ETTK)
Semantic Web	It includes Ontology, Knowledge-base Representation, Description Logic, and Agent Development for the next Generation Web - the Semantic Web. It is designed to use OWL, DAML+OIL, RDFs, RDF, or XML syntax to design ontology; developed using J2EE
Semantic Web Assistant	The Semantic Web Assistant combines the capabilities of production rule systems with RDF data on the Semantic Web. It lets users define rules that work with RDF data in order to carry out actions like e-mail notification etc.
SemanticWorks	A visual RDF/OWL Editor from Altova
Semantic Mediawiki	Semantic extension to the MediaWiiki wiki
Semantic Net Generator	Utility for generating topic maps automatically
SemWeb	SemWeb for .NET supports persistent storage in MySQL, Postgre, and Sqlite; has been tested with 10-50 million triples; supports SPARQL
Sesame	Sesame is an open source RDF database with support for RDF Schema inferencing and querying. It offers a large scale of tools to developers to leverage the power of RDF and RDF Schema
SHAME (Standardized Hyper Adaptible Metadata Editor)	SHAME is a metadata editing and presentation framework for RDF metadata. Annotation profiles are then used to generate user interfaces for either editing, presentation or querying purposes. The user interface may be realized in a web setting (both a jsp and velocity version exists) or in a stand alone application (a java/swing version exists)
SMART	System for Managing Applications based on RDF Technology
SMORE	OWL markup for HTML pages
SOFA	SOFA is a Java API for modeling ontologies and Knowledge Bases in ontology and Semantic Web applications. It provides a simple, abstract and language neutral ontology object model, inferencing mechanism and representation of the model with OWL, DAML+OIL and RDFS languages; from java.dev
Solvent	Solvent is a Firefox extension that helps you write Javascript screen scrapers for Piggy Bank
SPARQL	Query language for RDF
SPARQLer	SPARQL query demo and service
SPARQLette	A SPARQL demo query service
SPARQL JavaScript Library	SPARQL JavaScript Library interfaces to the SPARQL Protocol and interpret the return values as part of an Ajax framework
Surnia	Surnia can check an OWL ontology/knowledge base for inconsistency and entailments. It is implemented as a wrapper around first-order theorem prover

	(OTTER, for now at least). Unlike Hoolet (which turns the OWL into FOL), Surnia just turns the OWL into triples and mixes in axioms
SWCLOS	A semantic Web processor using Lisp
SWI-Prolog	SWI-Prolog is a comprehensive Prolog environment, which also includes an RDF Triple store. There is also a separate Prolog library to handle OWL
Swish	Swish is a framework for performing deductions in RDF. It has similar features to CWM. It is written for Haskell developers
Swoogle	A semantic Web search engine with 1.5 M resources
SWOOP	A lightweight ontology editor
Thema	Thema is an XML based data format (DTD) for thesauri, glossaries, lexicons, conceptual maps etc. up to ontologies. It contains publishing tools to convert into HTML, RDF etc. and to read different formats and is has a connection to the Semantic Web
Timeline	Timeline is a DHTML-based AJAXy widget for visualizing time-based events. It is like Google Maps for time-based information
TopBraid Composer	Top Quandrant's TopBraid Composer is a complete standards-based platform for developing, testing and maintaining Semantic Web applications
Trellis	Trellis is an interactive environment that allows users to add their observations, viewpoints, and conclusions as they analyze information by making semantic annotations to documents and other on-line resources
Tripple	TRIPLE is an RDF query, inference, and transformation language for the Semantic Web
Trippi	Trippi is a Java library providing a consistent, thread-safe access point for updating and querying a triplestore. It is similar in spirit to JDBC, but for RDF databases
Tucana Suite	Northrop Grumman's Tucana Suite is an industrial quality version of the Kowari metastore
Turtle	Terse RDF "Triple" language
W3C's RDF Validator	W3C's RDF Validator
WebCAT	WebCAT is an extensible tool to extract meta-data and generate RDF descriptions from existing Web documents. Implemented in Java, it provides a set of APIs (Application Programming Interfaces) that allow one to analyze text documents from the Web without having to write complicated parsers
WebOnto	WebOnto supports the browsing, creation and editing of ontologies through coarse grained and fine grained visualizations and direct manipulation.
Welkin	Welkin is a graph-based RDF visualizer.
WGFA	WGFA (Web Gateway for Fact Assessment) is a web application to create and manage W3C-OWL based ontologies, index websites, extract XML-RDF or Dublin-Core metadata and provide search and query operations on the websites based on the created semantic webs
Wilbur	Wilbur is lisp based toolkit for Semantic Web Programming. Wilbur is Nokia Research Center's toolkit for programming Semantic Web applications that use RDF written in Common Lisp
WOM	The IBM Web Ontology Manager (WOM) is a lightweight, J2EE Web-based system for managing Web Ontology Language (OWL) ontologies. It enables developers to browse or search the ontologies registered with the system by class or property names. In addition, they can submit a new ontology file
Wraf	Wraf (Web resource application framework) implements a RDF API that hopes to realize the Semantic Web. The framework uses RDF for data, user interface, modules and object methods. It uses interfaces to other sources in order to integrate all data in one environment, regardless of storage

WSMO Studio	A semantic Web service editor compliant with WSMO as a set of Eclipse plug-ins
WSMT Toolkit	The Web Service Modeling Toolkit (WSMT) is a collection of tools for use with the Web Service Modeling Ontology (WSMO), the Web Service Modeling Language (WSML) and the Web Service Execution Environment (WSMX)
WSMX	Execution environment for dynamic use of semantic Web services
xml2owl	Up to now, most ontologies are created manually, which is very time-expensive. The goal is it, to produce ontologies automatically via XSLT, which fit as good as possible to a given XML-file resp. XML-Schema-file
XML Army Knife	XML Army Knife
XMP	A labeling technology from Adobe that enables data about a file to be embedded as metadata into the file itself.
YARS	YARS (Yet Another RDF Store) is a data store for RDF in Java and allows for querying RDF based on a declarative query language, which offers a somewhat higher abstraction layer than the APIs of RDF toolkits such as Jena or Redland

Appendix E Acronyms

AFC2ISR	Air Force Command and Control, Intelligence, Surveillance, and Reconnaissance Center
AFRL	Air Force Research Laboratories
AFSAB	Air Force Scientific Advisory
BFT	Blue Force Tracking (COI)
C2-SSA	Command and Control Space Situation Awareness (COI)
CBA	Capabilities Based Assessment
CBO	Common Battlespace Object
CD&D	Content Discovery & Delivery
CDM	Core Data Model (OSD/NII)
CJCSI	Chairman Joint Chiefs of Staff Instruction
CMD	Common Mission Definition (schema)
COI	Community of Interest
CoT	Cursor on Target (program)
DAU	Defense Acquisition University
DCGS	Distributed Common Ground Station
DDMS	DoD Discovery Metadata Specification
DIA	Defense Intelligence Agency
DISA	Defense Information Systems Agency
DISR	Defense Information Standards Repository
DITSCAP	DoD Information Technology Security Certification and Accreditation Process
DoDIIS	DoD Intelligence and Information Systems
DOTMLPF	Doctrine, Organization, Training, Material, Leadership, Personnel, and Facilities
EO	Executive Order
ER	Entity Relationship (diagram)
ESC	(USAF) Electronic Systems Center
ETL	Extract, Transform, Load
FAA	Functional Area Analysis
FIOP-NBS	Family Of Interoperable Operating Pictures Network-Based Services
FNA	Functional Needs Analysis
FSA	Functional Solution Analysis
GIG	Global Information Grid
GIG-ES	Global Information Grid Enterprise Services
GML	Geography Markup Language
IA	Information Assurance
IC	Intelligence Community
IC-ISM	Intelligence Community Information Security Marking (schema)
IRTPA	Intelligence Reform and Terrorism Prevention Act
ISR	Intelligence, Surveillance, and Reconnaissance
JAMD	Joint Air and Missile Defense (COI)
JBI	Joint Battlespace Infosphere (program)
JC3IEDM	Joint Command, Control, and Communications Information Exchange Data Model
JCIDS	Joint Capabilities Integration and Development System
JCS/J2T	Joint Chiefs of Staff, Deputy Directorate for Targeting
JEFX	Joint Expeditionary Forces Experiment

JFCOM	US Joint Forces Command
JITC	Joint Interoperability Testing Command
JMBC2	Joint Battle Management Command and Control
JTM	Joint Track Management (COI, data model)
KB	Knowledge Base
KIP	Key Interface Profile
M2M	Machine to Machine Interoperability
MDA	1) Model-Driven Architecture 2) Maritime Domain Awareness (COI)
MDR	DoD Metadata Repository
NCDS	Network-Centric Data Strategy
NCES	Network-Centric Information Services (program)
NCOW	Network-Centric Operations and Warfare (reference model)
NII	Networks and Information Infrastructure as in OSD/NII
NR-KPP	Net-Ready Key Performance Parameter
NSS	National Security Systems
ODM	Ontology Definition Metamodel
ODNI	Office of the Director for National Intelligence
OGC	Open Geospatial Consortium
OIM	Operational Information Management
OJAWG	Operational Joint Architecture Working Group (data model)
OMG	Object Management Group
OSD	Office of the Secretary of Defense
OWL	Web Ontology Language
PIM	Platform Independent Model
PKI	Public Key Infrastructure
PSM	Platform Specific Models
QoS	Quality of Service
RDF	Resource Descriptor Framework
SDK	Software development kits
SI	Semantic Interoperability
SLA	Service Level Agreement
SME	Subject Matter Expert
S-O Mapping	Schema to Ontology Mapping
SOAF	Service Oriented Architecture Foundation
SOI	Scheme of Individuation
STRATCOM	US Strategic Command
SW	Semantic Web
TRL	Technology Readiness Levels
TST	Time Sensitive Targeting (as in TST COI)
USAF	US Air Force
VMF	Variable Message Format
W3C	World Wide Web Consortium
WSDL	Web Services Definition Language
WSMO	Web Service Modeling Ontology
XML	Extensible Markup Language
XSD	XML Schema Definition
XSWTP	Cross-Service Weapon Target Pairing