1.0 Introduction

The Data Access Function (DAF) provides net-centric services and means to access information within and relevant to the Warfighter Information Processing Cycle (WIPC). DAF services are invoked by all components of the WIPC, as shown in Figure 1-1. Within the Combat System, the DAF provides to sensor, track, reference, context, and sensor tasking and queuing information. The DAF consists of many data access services needed to meet the broad range of QoS, IA, and topology requirements and information types accessed across the WIPC. The DAF helps WIPC services operate autonomously with respect to each other; by separating the functionality of the service from the data, the services interact via the commonly understood and accessed data without any knowledge-of or explicit interaction-with each other.

The DAF supports WIPC information visibility, accessibility, understandability, and trustability across operational and security domains:

- Visible because DAF services are used to publish discovery metadata and then to query discovery registries.
- Accessible because DAF services are available to access all WIPC data using standard Plugs across all Topologies.
- Understandable because DAF services use a Common Core ontology with its foundational layer and support for Pedigree and Source Metadata.
- Trustable because DAF services enforce WIPC business rules for managing data.

The DAF supports WIPC data access across information Communities of Interest (COI) and Naval engineering and acquisition lines of business: subsurface, surface, air, land, and C4I.

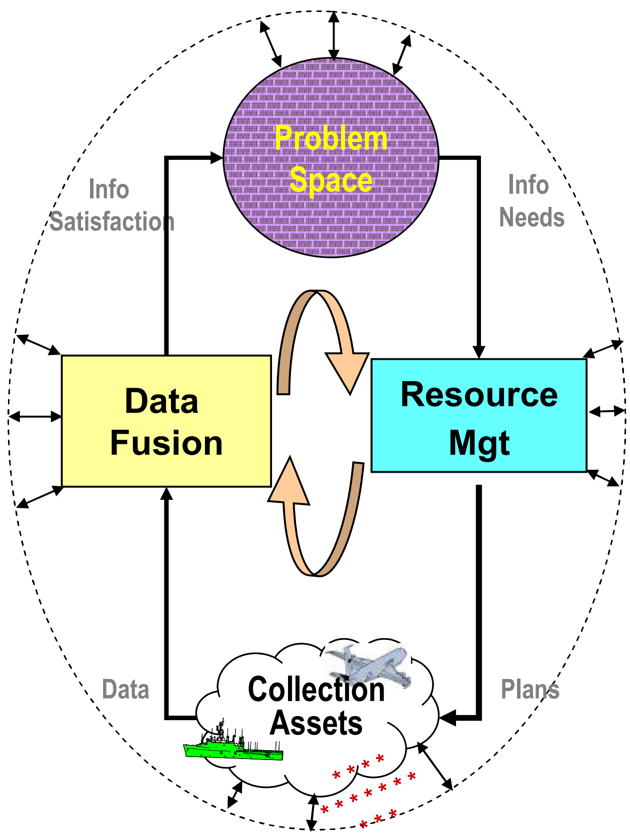

Figure 1-1. DAF Services Support All Components of the WIPC

The DAF is related to other WIPC concepts as shown in Figure 1-2. This figure shows an important type of Data Access Function is to retrieve P&SM data and, conversely, that P&SM data points to lineage (or descendancy) and Source Metadata objects that may need to be accessed. The Data Access Function provides the means to publish, query subtrees, and link P&SM data using GUIDs.

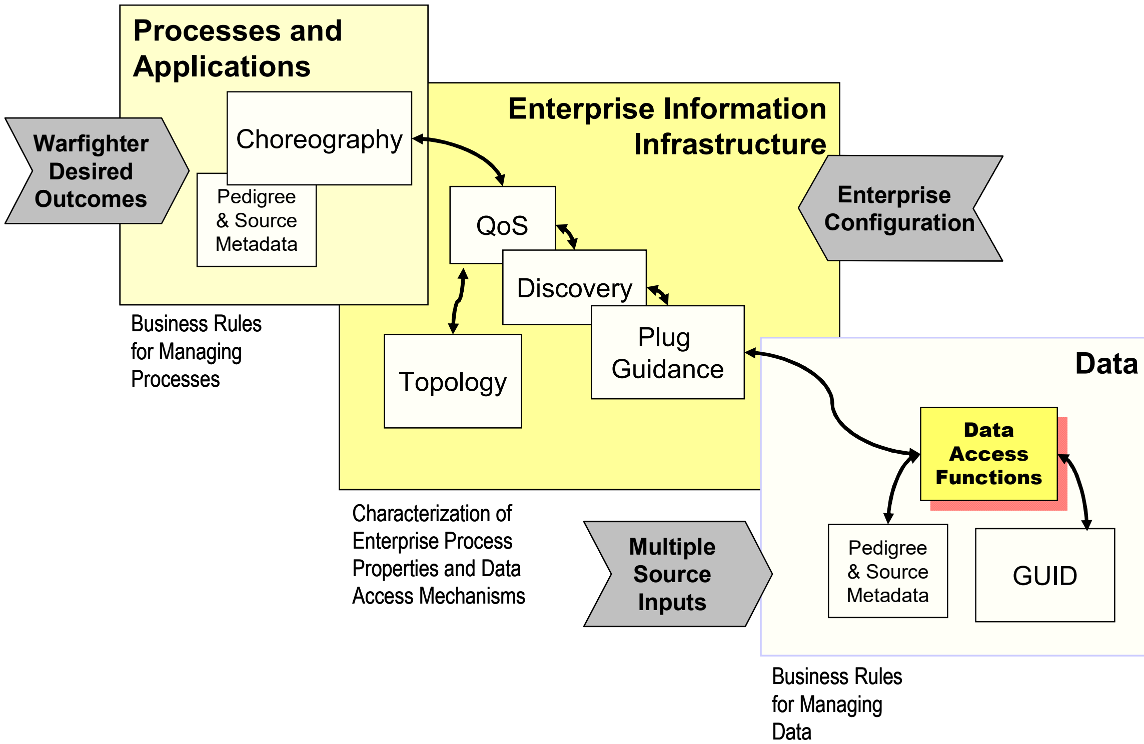

Figure 1-2. WIPC Concepts Relationships

The purpose of this document is to present a scoping and bounding definition of the Plug Guidance concept and its relationship to the other concepts that enable the IPC. This document will be followed by a specification document providing the actual Fusion Framework Interface Specification detailed guidance. This guidance will be in the form of specific requirement statements, best practice recommendations and examples (using the NESI guidance approach). Rationale supporting each specific guidance will be desirable. Guidance "Requirement" statements "must" be implemented and are compliance testable. Guidance "Best-practice" recommendation "should" be optionally implemented, depending upon the appropriate circumstances and resources.

2.0 Data Access Function Guidance Concept Discussion

The DAF encompasses many types of data access service functions. Access functions cover a range of atomic object retrieval, update, archive, and delete requests to more complex data graph operations. Access services also range over many types of subscription request, e.g., notification only, every-update delivery, conditional notifications and deliveries. Access services need to deal with a range of: (1) information and object types, (2), information and object semantics, (3) data QoS and IA requirements, (4) topologies, (5) data ID schemes, (6) physical format needs, and (6) business rules for information synchronization, collaboration, and exception handling over multiple nodes. The service function types along with these factors constitute a data services taxonomy. Note that while many of these might have been handled by a "gateway" in pre-SoA architectures, SoA leads to a de-bundling of these so they can be more flexibly choreographed, e.g., TADIL-J mediator and TN-GUID services might be choreographed with a Link-16 routing service in one case, but in a Tactical SoA bridge service in another. Each of these is discussed as follows.

a. Information and Object Types Accessed. The DAF needs to be able to support access to a wide range of object types such as Materiel (which includes Aircraft, Vessels, and Vehicles as well as consumable materiel), Organizations, Locations, Facilities, People, Geophysical Features (e.g., terrain elevations, meteorology), Control Features (e.g., shipping lanes), Tasks, (past, present, future, possible – own and other force), and Plans. Observation data can range from streaming video to human intelligence reports to real-time radar track reports. For any particular type of object, there are differing types of information. Examples are shown in Figure 2-1.

Object Type ➔	Materiel	Facility	Geophysical	Network	Organization	Social
	examples					
WIPC Info Category	SEAD aircraft / chemical gas	SAM Site	Huricane	Electrical	Terrorist	Holiday
Detects	radar amplitude / aerosal detector	image match	radar amplitude	power station	cell	news report
Features	Jet engine modulation / spectral features	site layout	diameter, shape	topology	size	duration, geo-extent
Kinematics	state & covariance estimate / cloud drift and dispersement	location	location, velocity, altitude	locations of nodes	past, present, predicted locations	key locations
Attributes	number ASM / gas type	radar and missile types and quantities	wind speed, precipitation rate	voltage, current	organization hierarchy	history, rationale
Activities	strike / village attack	operating air defense	building, dispersing	output levels, maintenance	attack planning	holiday events
Capabilities	range / lethality	coverage, max quantities	est. level	maximum power output, grid resiliance	max coordinated attack	max size, max population reach
Intent	SAM site neutralization / local population suppression	deny force A passage	target path	reach area	next target	population morale

Figure 2-1. DAF Services Operate Over Several Dimensions

1. Object Semantics. The data structure of fielded systems has proven to be an obstacle to successful transition of operationally valuable and needed technologies. This could be because they are closed (e.g., proprietary or difficult to gain access to) and/or because they are not modular or extensible so that any changes have costly impacts on other components of the system. However, it is unrealistic to expect that all services will employ a single standard or common semantics. So to be able to exchange data across semantic boundaries, DAF provides mediation (or translation, transformation) services. DAF employs a canonical model (schema) so the mediation will be semantically coherent, managerially tractable, and programmatically affordable.

b. Data QoS and IA Requirements. The DAF services accommodate the range of QoS and IA requirements as dictated by the mission. For example, in some cases lesser target accuracy may be a mission trade-off with timeliness of delivery.

c. Topology. In many cases data will originate or will be destined for legacy networks or systems, e.g., Link-16. In those cases, DAF services will employ router services to route data between the networks and their specific protocols. In other cases where independent SoA networks have been configured, e.g., tactical SoA employing DDS, a bridge service may be employed.

d. Data ID’s. A DAF service may require the use of a Data Identifier service to mange data identifiers between networks. For example, TADIL-J Track Numbers (TN’s) will need to be cross-referenced to GUIDs.

e. Physical Format Needs. A particular piece of data may be provided by DAF services in needed physical formats including XML and IDL.

f. Business Rules for Managing Data. Current data management is by fixed pre-engineered rules. In the WIPC, with non-specific and unintended users, the business rules are part of the service contract. This allows them to be more flexible. For example, TADIL-J’s rules regarding track management where only the source that reports is the one with the best position accuracy results in non-reporting of potentially value-added data. This is done for a variety of reasons such as communications bandwidth and end-processor capability limitations. In the WIPC, the DAF services could allow for data exchange tailored to more optimal estimation. Business rules can be specified in a variety of ways including Object Constraint Language (OCL) and RuleML.

3.0 IPC Operational Benefit

DAF operational benefits include:

a. Aid to Engagements by allowing visibility into and access and trusted use of amplifying or augmenting information that could aid target prioritization and scheduling, the prosecution of the target, (e.g., improved target designation, e.g., maneuver, counter-measures, cross-section), speed reaction time, collateral damage risk management (e.g., fratricide reduction), and improve accuracy and speed of kill assessment

b. Aid to Maneuver by providing access to more complete information for maneuver objectives as well as the maneuver environment"

c. Support to Command and Control understanding in that the DAF is the means for information sharing, thus SA and knowledge sharing. It supports development of knowledge because sharing information and knowledge is valuable to such.

d. Aid to Operational Planning analysis of alternative Courses of Action by unfettered access to needed information.

e. Support to Intelligence, Surveillance, and Reconnaissance Processing and Exploitation, which rely on access to large volumes of information and that will be aided by the uniform access.

f. Support to Intelligence, Surveillance, and Reconnaissance Processing and Exploitation, which rely on access to large volumes of information and that will be aided by the uniform access.

g. Also support to ISR dissemination in a complete and timely yet trustable manner.
h. Support to Situation Awareness with access to the volume and variety of data required for Situation Awareness.
i. Enabler for Battlespace Awareness Integration in that in order to integrate, must access.
j. Support to Enterprise Services like Information Sharing in that DAF is the means for information sharing.
k. Support to Position, Navigation, and timing data that can be accessed via DAF.
l. Support to Secure Information Exchange because DAF uniform and standard access services can be uniformly secured.
m. Aid to Protection through complete information access for knowledge of Contexts which may influence earlier threat detection as well as reduction in false alarms or collateral damage incidents."
n. Flexible QoS can more closely match operational mission requirements than fixed QoS.
o. Business Rules for data management than can be tailored or selected to meet the operational mission requirements.
p. A consist and uniform service form should reduce training costs and improve interoperability at the operator level.
q. Standard and common taxonomies used to describe data should also aid interoperability and reduce taxonomy translation costs.

4.0 Enterprise Environment Interactions

4.1 DAF in the Enterprise

At the Enterprise level, the DAF provides the following additional capabilities:

a. Enterprise-wide data sharing and accessibility with no obfuscation.
b. Enterprise-wide data understanding through use of the common ontology and standard / common object and information taxonomies.
c. Enterprise-wide independence of logical discovery and physical data access. DAF provides flexibility for information retrieval that gives a consumer options in the event that a data publisher provides multiple physical format DAFs for a single logical registry entry. DAF separates data access from physical location.
d. Track Servers can provide data to or receive data from the Enterprise.
e. Use of Enterprise Level 2 and 3 information for Level 0 signal processing (detection, feature extraction) and Level 1 Composite and Tactical System tracking.
f. Use of Enterprise Reference Data and Context for Level 0 signal processing (detection, feature extraction) and Level 1 Composite and Tactical System tracking. For example, knowledge of a nearby country's ballistic missile capabilities (level 2 information) may influence the identification of a ballistic missile seen (level 1 assertion). Another example might be knowledge of SAM sites and their capabilities and operational status

influencing detection (lowered detection threshold in site range), tracking (SAM trajectory type), and type estimation.

g. Use of Enterprise Authoritative sources instead of locally accessible convenience sources or Authoritative Source “surrogates”. For Authoritative Sources that change over time, either by added, deleted, extended, or changed data, Authoritative Source references will be more accurate.

h. “Gateway” access to legacy and external data sources through mediation, augmentation, bridge, and router services.

i. Exception and collaboration protocols (e.g., ID Conflict, Correlation disagreement, aberrancy removal / auto rebuild) at the Enterprise level so that more participants can be included in the estimation and adjudication. DAF protocols can be fixed or dynamic. The fixed protocols can be determined as part of a planning/setup process or drawn from a pre-determined setup. Dynamic protocols can be transitions between pre-determined protocols all the way to self regulating networks.

j. Loose coupling at the Enterprise level for services as components that performs a function and produce data. Data and algorithm component services are modular.

k. Uniform assertion and belief update structure, including Pedigree and Source Metadata across the Enterprise so that data across the Enterprise can be used to contribute to a object or situation understanding.

l. Responsive to QoS requirements

4.2 DAF Requirements on the Infrastructure

Infrastructure services must support

a. DAF services will need to register with enterprise for resources

b. Data Access Functions. The DAF provides services for various subscription types:

1. Auto send
2. Auto notify
3. One time immediate request (pull)
4. The DAF services address varying levels of persistence:
5. Immediate availability
6. Special request pull
7. Conditional notifications and deliveries
8. Update and delete requests, demands, etc.
9. Archive pull

c. DAF services support a broad range of data QoS such as:

1. Granularity (e.g., weather vs. grib fields)
2. Real-time short-latency “push” subscribers
3. Real-time longer latency notification subscribers

4. Discovery notifications
5. Short latency "pulls"
6. Longer latency queries
7. QoS for authitativeness, integrity, trustability, accuracy, etc.

d. Mediation / Transformation. The canonical model is the common ontology which has many levels of abstraction. Because of this abstraction, source data elements can be either, (1) mapped to a more general data element and then translated to a destination in a "hub and spoke" manner, or (2) treated as a COI specialization. It is important to note, however:

1. Even if the mapping from the source "spoke" to Common Core "hub" is simple and the mapping from the Common Core "hub" to the destination "spoke" is simple, the translation from the source semantic form to the destination will typically be complex, i.e., many-to-many. It may also be context dependent. Context dependent translation requires "state" knowledge, either by persistence in the mediator or access to some track server or persistent data service. When possible, means should be engineered to avoid context dependent translation.

e. Semantic Augmentation. In many cases, the semantic scope of a source may not include that of a request, e.g., a TADIL-J source will not have the semantic scope for Common Core uncertainties, pedigree, and sensor metadata. In these cases, the mediator may need to employ an "augmentation" service that can estimate the missing information. For example, knowing the source ship or aircraft for the TADIL-J report, it is typical to estimate the uncertainties and sensor metadata based on the principal sensor for the type of TADIL-J report.

f. Discovery Metadata. The discovery metadata will need to describe the object type and information category using understandable (e.g., standard or common) taxonomies, specific object attributes about the information (e.g., where, when, immutability, versions available) along with the DAF service(s) available such as their data function(s), mediations, data QoS and IA attributes, topology handling, data ID services, and business rules.

g. Discovery Services and Metadata Registry. Discovery services will access metadata registries to access to changing Authoritative sources and provide dynamic, runtime binding to publishers IAW QoS, IA, and topology requirements. For example, the nearest Authoritative source may be chosen over a more accurate Authoritative source if timeliness requirements so dictate.

h. Support for Standard or common taxonomies for discovery. These can be structured in super-subtype taxonomies or more generalized to ontologies with whole-part and other relationship types. There is much source material for these such as Mil-Std 6016, Mil-Std 2525, and JC3IEDM. Many of these were collected and partially integrated as part of Navy OAFn experimentation with a Common Data Model.

i. Metadata Registration. Some of registry descriptive elements will be requirements on a publisher. Mediation, Augmentation, Routing, and/or Bridge services may need to provide these for legacy or external sources of data. For example, a TADIL-J mediation service will need to add TADIL-J metadata tag names and descriptions.

j. IA. Assumes black Core (MLS)

1. Debundle classification levels
2. Authentification / security IO – will use services provided by Enterprise

k. Data Management Business Rules. Consequently, the DAF services will span a broad range, for example:

1. Master synchronizations
2. Collaborative synchronizations

l. Routing and Bridging. In many cases, DAF services will employ services of a capability provider, Program of Record, or other external program wherein DAF provides the common Common Core data layer (mediation from POR) and Plug so that WIPC services can be choreographed without the need to consider mediations and multiple Plug configurations.

m. Archiving. TBS

n. Persistent Storage. TBS

5.0 Challenges

Unresolved topic(s):

a. Mapping and mediation (translation, transformation) are difficult and costly throughout the IT community despite many developments for heterogeneous data interoperability, data layering, canonical models / schemas, ontologies / taxonomies, data warehouse Extraction, Transformation, and Loading (ETL) tools, etc. Many of the available tools operate more at the physical syntax level than at the difficult semantic level.

b. Semantic Augmentation. Particularly to deal with unavailability of track state, estimates, error states.

c. Undo / Aberrancy Removal. TBS

d. Exchange Format. Although XML is does not require foreknowledge of schema and pre-runtime arrangement to receive such data, as compared to binary (e.g., IDL), it may not meet all WIPC QoS requirements (e.g., bandwidth to submarines, real-time exchange between cooperative tracking radars).

e. Taxonomy Standardization / Commonization. Future interoperable data exchange is dependent on establishing core and COI taxonomies and the rules for translating across dialects. There are several challenges here:

1) The importance to interoperability is underappreciated because they are encoded in the software as deep details. Yet they often contain some of the more significant operational semantics.
2) Often existing “taxonomies” are in the form of flat lists, e.g., the JC3IEDM “category codes”. A SME can guess at an intended structure, but it is of mixed type. Often they are undefined, e.g., TADIL-J.
3) Modeling languages and tools poorly support structure. For example, Entity-Relationship models do not have a model feature for structured domain values.

Similarly, UML Class Models do not have model feature for structured enumeration values. This has been barely noted in the literature, e.g., a literature search turns up few instances such as [1].

4) There are many sources for these that typically mix taxonomic (or ontologic) relationship types, e.g., super-subtype, whole-part. Often multiple typing rules are used. The mixing is non-explicit (undocumented) and must be analyzed-out in order for the taxonomies to be combined and for them to work as required. Tools to support such analysis and combination (e.g., Protégé [2], Top Braid [3], Navy DIAD [4]) are in their infancy and insufficient.

f. Archive Requirements Analysis.

g. Object - Relational Interoperability. The technologies for object and relational overlay exist but are awkward and customized.

h. Real-time Persistent Storage and Synchronization. Persistent storage is either customized, employs limited market OODBMS', or relational DBMS' that does not map well to real-time class models. Master-slave or other market synchronization schemes have not been proven sufficiently in a Combat System scale environment.

i. Ability to backtrack track updates and changes and the quality of those

The next paragraph delineates the dependencies and assumptions that address some/all of these challenges.

6.0 Dependencies and Assumptions

Dependencies/interdependencies/interactions of Data Access Function with other IPC concepts are shown in Figure 6-1.

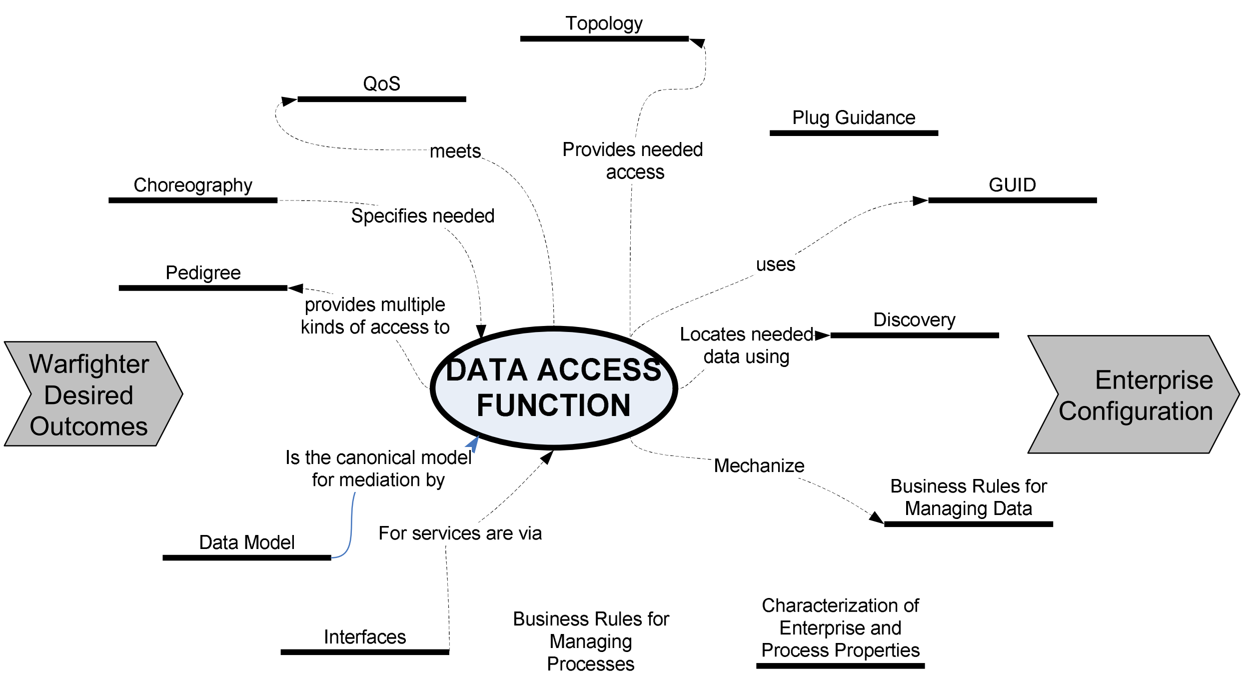

Figure 6-1. The DAF Fits with Other WIPC Concepts

More explicitly, DAF depends on

a. Data QoS that is complete with respect to operational mission needs and that is reliable in operation, that is, has been certified

b. Complete and current topology information and mechanisms for traversing the topology.

c. Being able to included in and utilize choreography services

d. The Common Core data model and the standard / common taxonomies for discovery

e. Pedigree to be able to access object source data, via the pedigree "chain" or "tree" if necessary

f. Discovery services that publicize the full range of data access attributes so the best source and be determined

g. GUID's that support complete and reliable access to objects or data fragments (graphs)

h. Interoperating data sources and destinations: (1) Track Servers, (2) Observation publishers, (3) legacy and external.

Assumptions, actions, conventions that could resolve, avoid, begin mitigation, or begin resolution, for these dependencies are:

a. Analysis of QoS requirements from an operational perspective for warfighting functions WIPC supports. Development of a QoS certification process.

b. Cross-network topology specification for selected legacy network, e.g., Link-16. Then an LTE.

c. Choreograph of a end-to-end data access process, first as an end-to-end service specification and then in an LTE.

d. Common Core data model formal methodology establishment and model review and update thereupon. Some emerging ontology methodology work may contribute in this area.

e. Taxonomy requirements analysis and development methodology development. Taxonomy tool development. Some on-going ONR work may contribute in this area.

f. Pedigree "chaining" experiment.

g. Experiment to extend NCES discovery services with needed attributes

h. GUID experiment with complex objects

i. Track Server experiment

Appendix A. Glossary

Add/update/delete items in this glossary to make it applicable to this concept paper. The glossary contains a concise definition of terms used within this document, but the full description in the text is the normative description.

Capability [JCIDS]
The ability to achieve a desired effect under specified standards and conditions through combinations of means and ways to perform a set of tasks. It is defined by an operational user and expressed in broad operational terms in the format of a joint or initial capabilities document or a joint doctrine, organization, training, materiel, leadership and education, personnel, and facilities (DOTMLPF) change recommendation.

Capability [OASIS]
A real-world effect that a service provider is able to provide to a service consumer.

Framework
A set of assumptions, concepts, values, and practices that constitutes a way of viewing the current environment.

Information Model
The characterization of the information that is associated with the use of a service.

Interaction
The activity involved in making using of a capability offered, usually across an ownership boundary, in order to achieve a particular desired real-world effect.

Pattern
A repeatable general solution to a commonly occurring problem. It is a combination of implicit and explicit knowledge repeatedly applied with success in the past and commonly captured as best practices and models.

Policy
A statement of obligations, constraints or other conditions of use of an owned entity as defined by a participant.

Process Model
The characterization of the temporal relationships between and temporal properties of actions and events associated with interacting with the service.

Quality
A general term applicable to any trait or characteristic whether individual or generic; a peculiar and essential character, an inherent feature, a distinguishing attribute, or an intelligible feature by which a thing may be identified.

Real world effect
The actual result of using a service, rather than merely the capability offered by a service provider.

Reference Architecture
A reference architecture is an architectural design pattern that indicates how an abstract set of mechanisms and relationships realizes a predetermined set of requirements.

Reference Model
A reference model is an abstract framework for understanding significant relationships among the entities of some environment that enables the development of specific architectures using consistent standards or specifications supporting that environment. A reference model consists of a minimal set of unifying concepts, axioms and relationships within a particular problem domain, and is independent of specific standards, technologies, implementations, or other concrete details.

Semantics
A conceptualization of the implied meaning of information, that requires words and/or symbols within a usage context.

Service
The means by which the needs of a consumer are brought together with the capabilities of a provider.

Service Consumer
An entity which seeks to satisfy a particular need through the use capabilities offered by means of a service.

Service Description
The information needed in order to use, or consider using, a service.

Service Interface
The means by which the underlying capabilities of a service are accessed.

Service Oriented Architecture (SOA)
Service Oriented Architecture is a paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with and use capabilities to produce desired effects consistent with measurable preconditions and expectations.

Service Provider
An entity (person or organization) that offers the use of capabilities by means of a service.

Software Architecture
The structure or structures of an information system consisting of entities and their externally visible properties, and the relationships among them.

Solution Space
A set of potential implementations all of which exhibit the architectural qualities expressed by an architecture description. This set of potential implementations becomes the set of candidate implementations evaluated during engineering development and from which a best implementation is selected through such development.

Appendix B. References

1 Sally Rice. John F. Roddick, and Denise de Vries; "Defining and Implementing Domains with Multiple Types using Mesodata Modelling Techniques", Conferences in Research and Practice in Information Technology, Vol. 53. Sven Hartmann, Markus Stumptner and Yasushi Kiyoki, Ed., Australian Computer Society, Inc., 2006

2 http://protege.stanford.edu/

3 http://www.topquadrant.com/topbraid/composer/index.html

4 NESI and NECC URLs to CDV

Mary K. Pulvermacher, "Implementing a Net-Centric Data Access Service", The MITRE Corporation, December 2007

Boris Lublinsky, "Incorporating Enterprise Data into SOA", http://www.infoq.com, C4Media Inc., November, 2006

Michael J. Carey, "SOA What?", IEEE Computer, IEEE, March 2008

Derek Henninger, "SOA and persistent data requirements", http://www.itarchitect.co.uk/, Bearpark Publishing, London, April 2005

"Service Oriented Architecture and the Army Service Layer", http://data.army.mil/datastrategy_serviceorientedarchitecture.html, U.S. Army CIO/G-6

"Army Data Services Layer Capabilities (ADSL)", http://data.army.mil/ADSL.html, U.S. Army CIO/G-6

"Service Data Objects White Paper", http://www.osoa.org/display/Main/SDO+Resources, Open SOA Collaboration, March 2007

Pat Helland, "Data on the Outside vs. Data on the Inside", http://msdn.microsoft.com/en-us/library/ms954587.aspx, Microsoft Corporation, 2008