Pedigree in the Warfigher Information Processing Cycle (WIPC)

All information in the WIPC has a source -- even a lineage of sources. Within the WIPC, information lineage is referred to as "Pedigree[1]" and information about the source is called, "Source Metadata (P&SM). Pedigree is a chain of observations or object beliefs[2] and along with a description of how such observations or object beliefs were arrived at while Source Metadata is a characterization of the source, whether it be a sensor, individual operators, or a system of machines and operators. P&SM lineage describes how a piece of information came about; P&SM descendancy describes how a piece of information was used. P&SM lineage is used to:

a. Assess the trustworthiness or quality of provided information by pulling the lineage chain and assessing the trustworthiness of the sources.

b. Augment the quality of provided information. Although certainty estimates (e.g., confidences, covariances) may be published or available for information, they may not tell the whole story when estimation model assumptions are violated.

c. Aid in fusing the provided information with other information

d. Support corroboration analysis[3] and avoid of information double counting[4] (also known as "data incest", "rumor propagation", or "data ringing")

P&SM descendants can be used to:

a. Understand how published information is used by others

b. Remove aberrancies (own-force "mistakes") or deceptions (opposing force)

c. Assess security vulnerabilities from inadvertently disclosed information

Within the context of the overall WIPC, the P&SM accompanies and explains Observations, Estimates, and Assessments as well as Needs, Tasks, and Plans, as shown in Figure 1-1. This figure shows that Pedigree chains (actually "trees") trail all Observations, Data Fusion outputs, Information Needs, and Collection Plans.

---

[1] "Provenance" is not distinguished from "pedigree" herein since the chaining of pedigree's constitutes provenance.

[2] From the perspective of P&SM in the IPC, all C2ISRT data is viewed as an assertion or belief.

[3] Marisa M. Gioioso*, S. Daryl McCullough, Jennifer P. Cormier, Carla Marceau, Robert A. Joyce; "Pedigree Management and Assessment in a Net-centric Environment"; Defense Transformation and Net-Centric Systems 2007; Proc. of SPIE Vol. 6578

[4] Pablo O. Arambel; "Structured Pedigree Information for Distributed Fusion Systems"; Signal Processing, Sensor Fusion, and Target Recognition XVII; Proc. of SPIE Vol. 6968

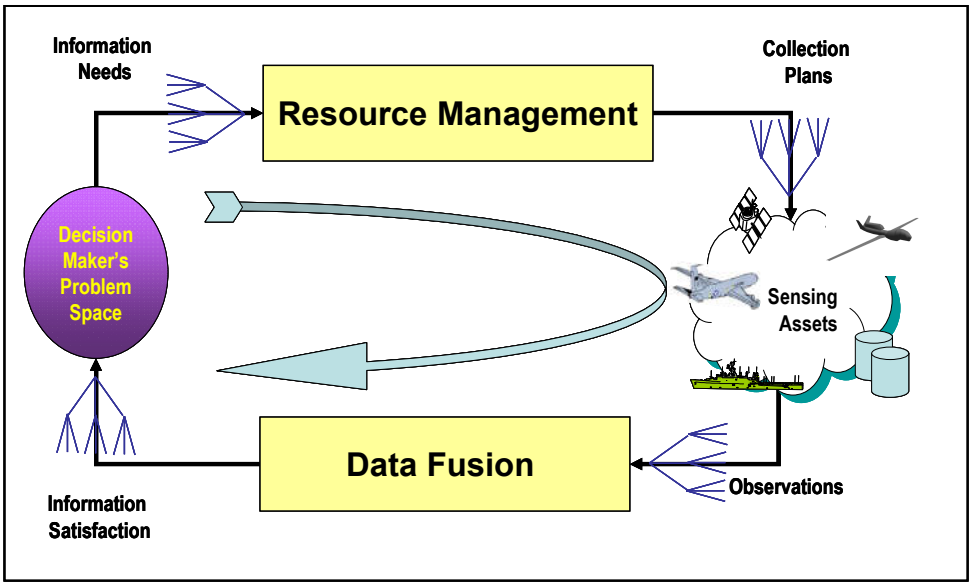

Figure 1-1. Pedigree "Trees" Exist for all WIPC Information

P&SM 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 Pedigree data points to lineage 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. .

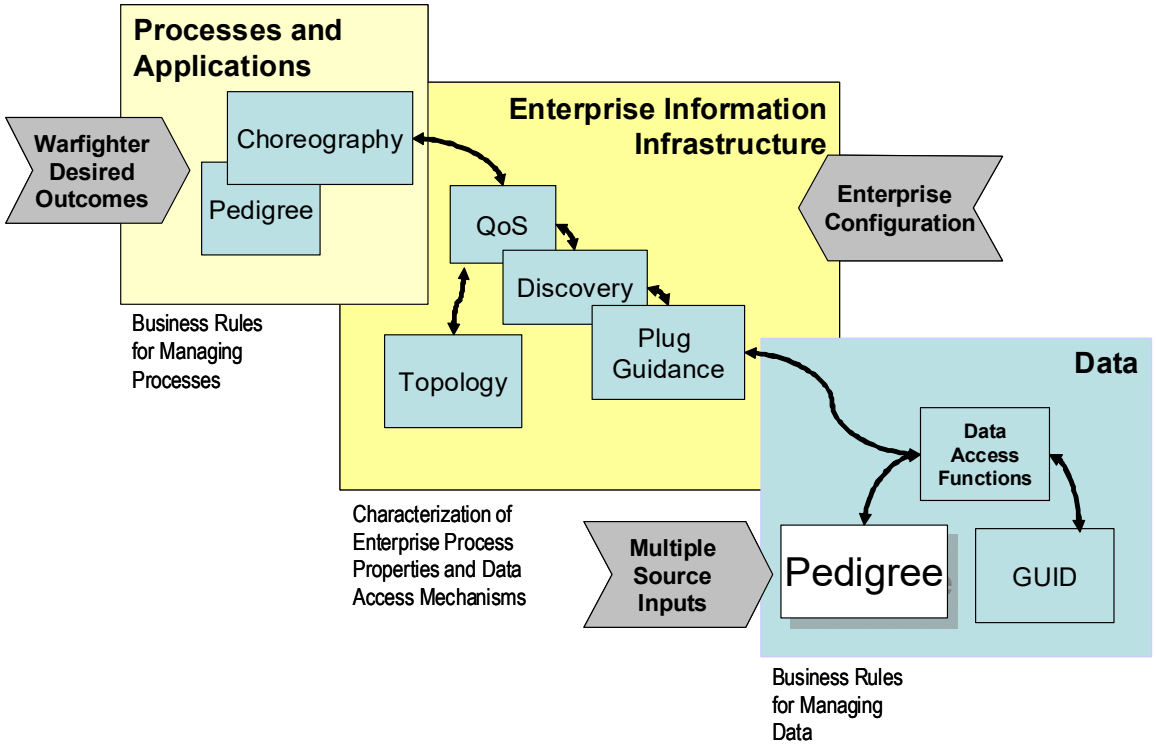

Figure 1-2. WIPC Concepts Relationships

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.

1.0 P&SM Guidance Concept Discussion

There are five key concepts to P&SM. First, P&SM consists of three major components as shown in Figure 1-1.

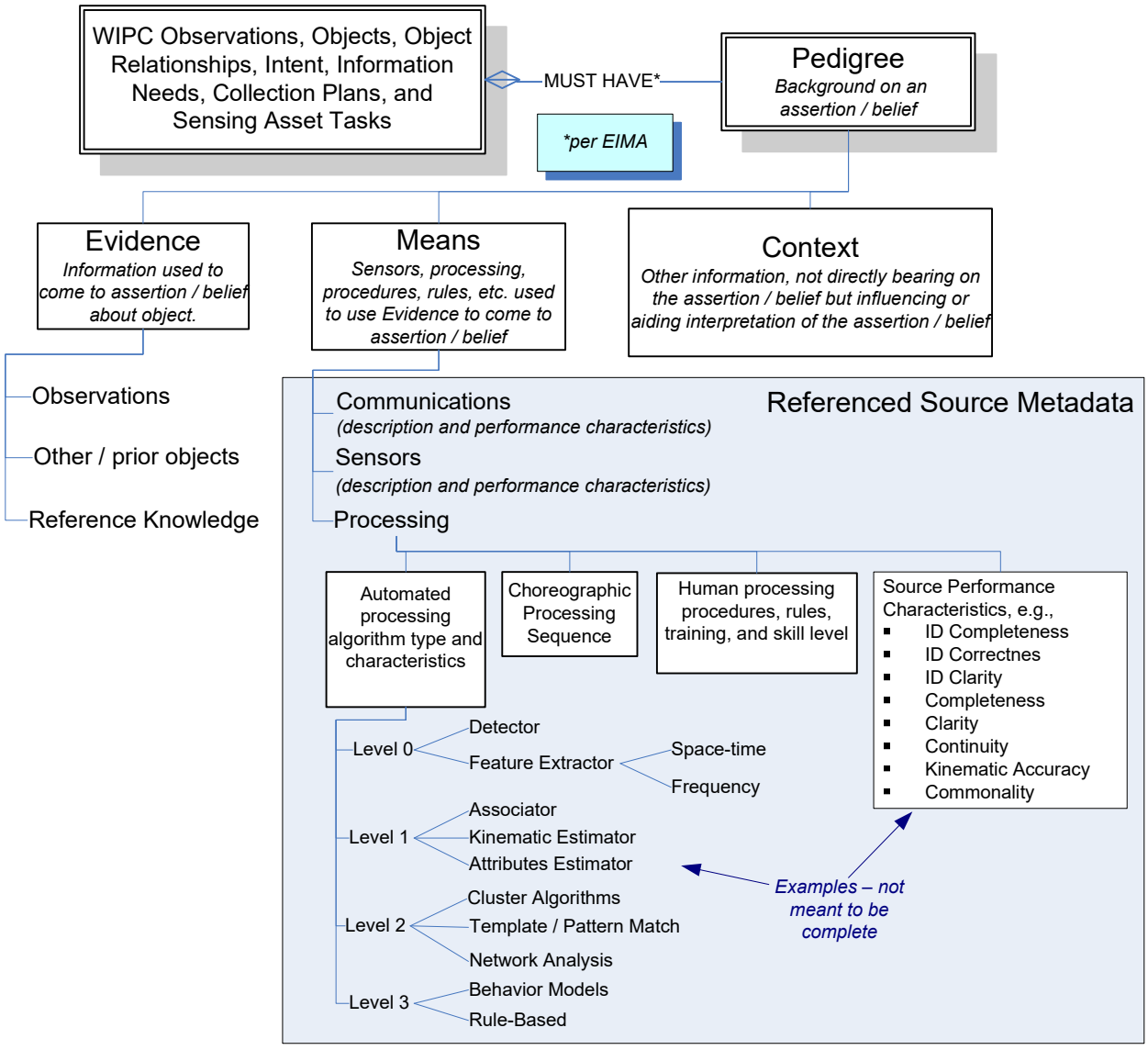

Figure 1-1. P&SM Components

Note that P&SM can be associated with the overall object as well as individual assertions / beliefs about the object. This is because in a multi-source environment, different information about the object may be, indeed, usually is, derived from different sources. For example, an aircraft's kinematics may be derived from a 3-D air search and tracking radar while the identity is derived from an ESM or IFF sensor.

The second key concept regarding P&SM is that it is logically structured as a line (or tree) of descent. It can be thought of as "chained" and dynamic. That is, the evidence used by a fusion node at one time can consist of assertions / beliefs made by another (or the same) fusion node at a prior time. This is illustrated in Figure 1-2.

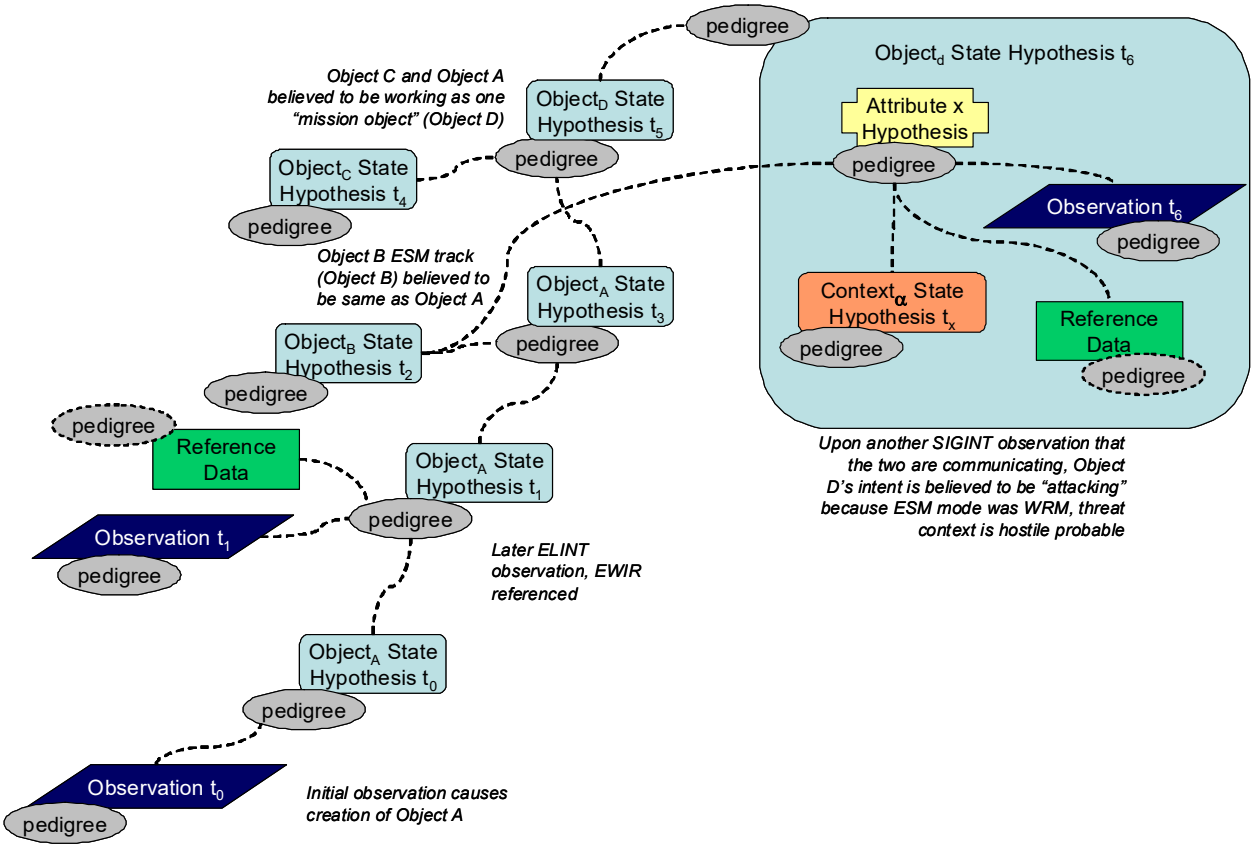

Figure 1-2. P&SM Chaining and Dynamic State Change

The third key concept regarding P&SM is Context. Context is relative, in what is regarded as Context is a function of the object in focus. For example, for aircraft tracking, weather can be considered Context but for weather prediction, a WIPC fusion process, the weather is the object of focus and Context might include Climatology, solar activity, known volcanic eruptions, etc. So while Context is treated a ground for certain WIPC fusion processes, it itself is the result of a WIPC fusion process. That is, there is a Context estimation process that is part of the WIPC. Note also, that Source Metadata can be considered a type of Context, that is, the Context of the observer or source. This differs from what may be considered "environmental" context, that is, the environment in which the object of interest is being conjectured about.

The forth key concept to P&SM is that it is drawn (or "pulled") only when needed. In most cases, it is not necessary to inquire as to the reasons for fusion node's assertions or beliefs. In addition, there may be different levels of detail or granularity of P&SM that would be pulled for different purposes.

Fifth, source Metadata must allow for information sharing while protecting sources and methods.

2.0 IPC Operational Benefit

P&SM's operational benefits range from Force Protection and Engagement and Maneuver and Operational Planning to Battlespace Awareness and Command and Control.

a. Being able to pull P&SM information on a target could provide quicker confidence that a target should be engaged prior to the engagement, somewhat analogous to the Mode 4 pre-engagement interrogation. After the engagement, it may be useful to pull P&SM information in conducting kill assessment using multiple sources of information of disparate types.

b. It may be valuable in formulating and evaluating alternative maneuver Courses of Action, being able to reach back into critical pieces of Situation Awareness data whose accuracy and interpretation could "swing" a CoA decision. Knowing the trustworthiness and accuracy of information could alter the maneuver commander of situations to pay particular attention to once the maneuver is underway and in helping pre-formulate risk mitigation alternative CoAs.

c. Resource Management. P&SM shows what sources have already be employed and that, therefore, may not have much additional value in re-tasking, thereby avoiding wasted resource utilization.

d. Integration of multiple sources may require P&SM for reconciliation of differences in assertions / beliefs between fusion nodes. For example, ID Conflict or correlation difference in the TADIL's. These are currently worked off manually via voice circuit. The inefficiencies and dis-interoperabilities inherent in this design have been shown in many Joint exercises.

e. Interpretation of an assertion or belief.

f. Determination as to whether the assertion or belief is, in fact, new information

g. Removal of aberrancies. The P&SM chain may have to be pulled to re-estimate the object.

h. Being able to pull the lineage and source information aids understanding of the information so fusion nodes can better collaborate in developing and sharing knowledge to achieve situation awareness.

Additionally, P&SM aids Information Assurance, e.g., by maintaining traceability for understanding who touched what data in support of vulnerability analysis

3.0 Enterprise Environment Interactions

3.1 P&SM in the Enterprise

The differences in P&SM in an Enterprise environment are as follows:

a. Before components / services were decoupled, the

b. Not used in past

c. All information has P&SM that is visible, accessible, and understandable by all services in the Enterprise.

d. P&SM is part of data's explicit, dynamic, and run-time describe-ability. Pre-Enterprise P&SM was implicit, e.g., via engineering design. Consequently, it was fixed at design time and was static.

e. Because its structure and means of access are standard across the Enterprise, it supports decoupling of components and modularity of algorithms components

3.2 P&SM Requirements on the Infrastructure

Infrastructure services must support

a. P&SM is normally pulled but can be subscribed to. For pulls or subscriptions, the subscriber will need to specify:

   1) The depth of chaining to be pulled / subscribed-to

   2) The components of P&SM required

   3) Where in the P&SM chain the query needs to continue forward from

b. P&SM topics that vary by granularity

c. Chaining by GOID / GUID

d. How to specify that last chain and what belief or assertion is being pulled

e. Cascading of P&SM choreography

f. Drill down confidence measures

g. Broker out of sequence, non-synchronous – choice or go right to

h. Source Metadata must have variable levels of IA granularity to protect sources and methods

4.0 Challenges

a. Specificity on P&SM pulls, especially near the tactical edge

b. Understandability of domain P&SM

c. Deconflict P&SM vis a vis "Sensor metadata" issues

d. Continuum of "P&SM" requirements

e. Sensor System requirements

f. Legal aspects

g. Sanitization is an issue (protecting sources and means)

h. Determining the scope and structure of P&SM and Source Metadata.

i. Source Metadata models

j. Volume of Pedigree retention vis-à-vis history and logging.

k. Inability to de-aggregate fused information

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

5.0 Dependencies and Assumptions

Within the context of the other WIPC concepts, P&SM fits as shown in **Error! Reference source not found.**:

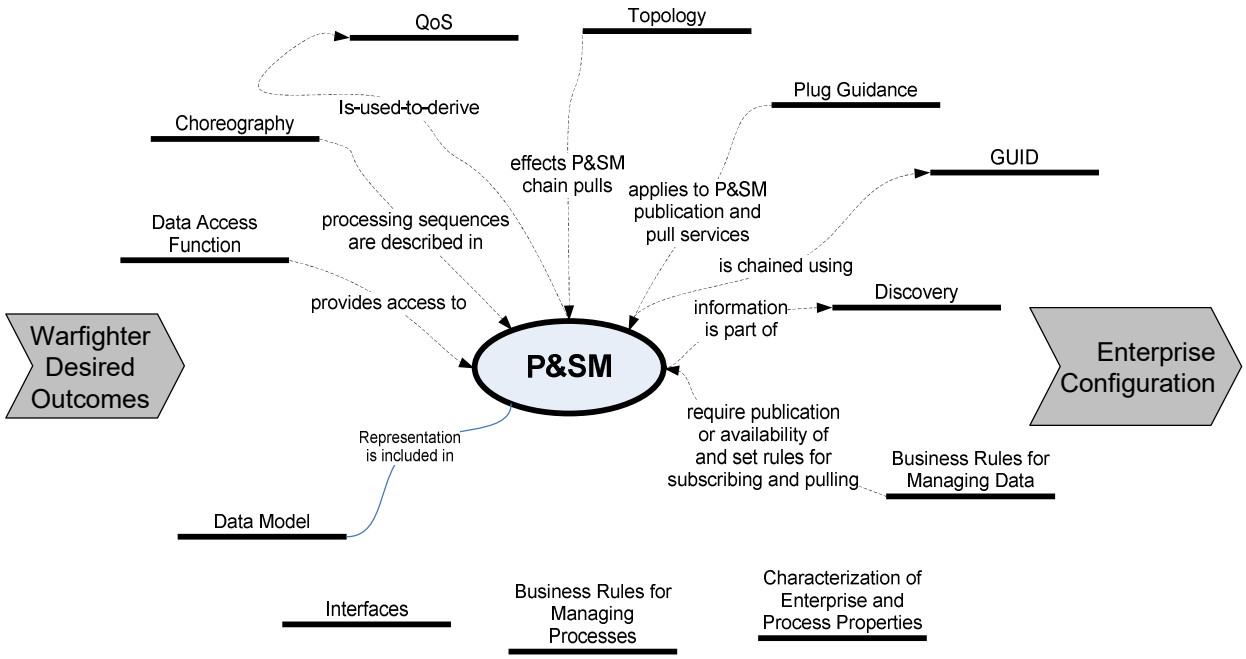

Figure 5-1. P&SM Concept Relationship with Other WIPC Concepts

More explicitly, P&SM depends on

a. publishers publishing or having protocol for requesting P&SM
b. timely response to P&SM requests, e.g., for pre-engagement confirmation
c. standard reference for sources
d. standard models / characterizations of processing, actions, etc. taken by fusion nodes in deriving the information
e. standard models for Source Metadata such as sensor models
f. functional protocols for fusion node conflict resolution after P&SM data has be pulled and differences still persist

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

a. An “augmentation” function within the Data Access Function that constructs, perhaps partial or approximate, P&SM, by knowledge of the publisher
b. Doctrine that engagement candidates are published with n-levels of provenance
c. GUID
d. Common Core extensions by fusion processing / sensor observation sub-COI

e. Leverage SensorML, TransducerML, and other COIs developing sensor models

f. ONR, AFRL, and ARL 62/63 research

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

[CJCSI6212]
CJCSI 6212.01D, Interoperability and Supportability of Information Technology and National Security Systems, 8 March 2006.

[DODAF]
U.S. Department of Defense, DoD Architecture Framework v1.5, 23 April 2007

[DOD5000]
U.S. Department of Defense, DODI 5000.2, Operation of the Defense Acquisition System, 12 May 2003.

[DOD4630]
U.S. Department of Defense, DODI 4630.8, Procedures for Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS), June 30, 2004

[FN SDF]
U.S. Navy, Naval Space and Warfare Systems Command, FORCEnet Service Description Framework v2.0, March 2007.

[JCIDS]
CJCSI 3170.01E, Joint Capabilities Integration and Development System, 11 May 2005.

[MERCER]
Mercer, B. D, "Thoughts on DoD Architecting and How to Improve the Practice v3.0." The MITRE Corporation, 12 June 2007.

[NESI]
U.S Navy, PEO C4I. Net-Centric Enterprise Solutions for Interoperability (NESI), 3 June 2005.

[NCES]
The MITRE Corporation, Net-Centric Enterprise Services (NCES) Reference Design, MTR 05W0000088, November 2005.

[OASIS]
Organization for the Advancement of Structured Information Standards (OASIS), "OASIS Reference Model for Service Oriented Architecture v1.0, Committee Specification", http://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf, 2 August 2006

[SEMP]
U.S Navy, PEO C4I, PMW 160, Systems Engineering Master Plan v2.0, 30 March 2007.

[SHAFFER]
Shaffer, Gary, Enterprise Services Primer, U.S Navy, PEO C4I, PMW 160, San Diego, California, December 2006.

[W3C WSA]
W3C Working Group Note "Web Services Architecture", http://www.w3.org/TR/ws-arch/, 11 February 2004