Department of Defense
Data Services Environment

Concept of Operations (CONOPS)
August 2, 2013

Novetta Solutions
(formerly FGM, Inc.)
12021 Sunset Hills Road
Suite 400
Reston, VA 20190

RECORD OF REVISIONS

REVISION DESCRIPTION	MODIFIED ON	MODIFIED BY
Removed references to ESM.	10/14/2014	Laura Butler

TABLE OF CONTENTS

1. MISSION AND SCOPE 1

2. BACKGROUND 1

2.1 DoD and Net-Centricity 1
2.2 Service-Oriented Architectures and the Cloud 2
2.3 Registry Federation 4
2.4 Document Organization 5

3. CONSTRUCT 6

4. CORE FUNCTIONS 6

4.1 DSE Capability Overview 6
4.1.1 User Dashboard Overview 9
4.1.2 Find Metadata Overview 11
4.1.2.1. Metadata Search and Discovery 11
4.1.2.2. Access, Retrieval and Subscription 11
4.1.2.3. Understanding Metadata 11
4.1.2.4. Identifying Metadata Linkages 11
4.1.3 Publish Metadata Overview 12
4.1.3.1. Publication (Submission) 13
4.1.3.2. Schema Validation 13
4.1.4 Reports Overview 13
4.1.5 Communities of Interest Overview 13
4.1.6 Community Overview 14
4.1.7 Help Overview 14
4.1.8 Feedback and Support 14
4.2 Descriptive Metadata about Authoritative Data Sources 14
4.3 Descriptive Metadata about Services and Capabilities 16
4.3.1 Benefits of DSE's Publication and Discovery Capabilities 17
4.3.2 Service Publication 19
4.3.3 Service and Capability Information Model Concepts 21
4.3.4 Concept Realization 23
4.3.5 Categorization 24
4.3.5.1. Modeling Service Taxonomies 25
4.3.5.2. Predefined Enterprise Taxonomies 25
4.3.5.3. Categorization of Taxonomies 25
4.3.6 Versioning 27
4.3.7 Lifecycle Management 27
4.4 Registry Administration 27
4.5 Interaction with Other CES and Structural Registries 29
4.6 Operational Use Cases 29
4.6.1 Discover Assets 31
4.6.1.1. Service and capability discovery 32
4.6.1.2. Static (Design-Time) Service Discovery 33
4.6.1.3. Dynamic (Run-Time) Service Discovery 33
4.6.1.4. Subscription Based Service Discovery 35
4.6.2 Govern 35
4.6.3 Discover Enterprise Metadata 36
4.6.4 Publish Authoritative Data Source Metadata 36
4.6.5 Publish Service Information 38
4.6.5.1. Taxonomy Management 38
4.6.6 Runtime (Machine-to-Machine Interactions) 39

5. SUMMARY 40

APPENDIX A: ACRONYM LIST 42

APPENDIX B: REFERENCES.....44

LIST OF FIGURES

Figure 1: Service-Oriented Architecture 2
Figure 2. Registry Federation 5
Figure 3. DSE Functional Overview 7
Figure 4. DSE 2.2 Home Capability 10
Figure 5. ADS - Core Concepts 15
Figure 6. Process of Establishing an Authoritative Data Source 16
Figure 7. Dynamic Service Discovery and Composition 19
Figure 8. Stages of Service Information Entity Lifecycle 20
Figure 9. Information Model Relationships 23
Figure 10. Abstract and Concrete WSDL Usage Scenario 24
Figure 11. Top-Level DSE Use Cases 30
Figure 12. Publish Data Artifacts 31
Figure 13. Discover Assets Workflow 32
Figure 14. Govern DSE and Metadata Resources 35
Figure 15. Federated or Enterprise Search 36
Figure 16. Authoritative Data Source Registration, Proposal, and Approval Process 37
Figure 17. Interactive Service Publication 38
Figure 18. Runtime Use Cases 39

1. Mission and Scope

The Data Services Environment (DSE) utilizes principles of service-orientation, and facilitates information sharing across the DoD, integrating with SOA-based services and enterprise services within the Program Executive Office Enterprise Services Directorate (PEO-ESD). DSE simplifies the publication and discovery of data services, and can be a key enabler for integration, search, discovery, interoperability, and governance of apps, services, and components critical to the DoD cloud environment of the future. The DSE revolves around the goals of DoD Net-Centric Data Strategy, requiring that data be made visible, accessible, understandable, trustworthy, interoperable, and responsive to user needs on a continuing basis.

2. Background

2.1 DoD and Net-Centricity

Net-Centricity is a concept of connected computing in which computers and computing are defined in the context of ubiquitous connectivity. Net-Centricity provides shared awareness and information among users in a network and allows real-time collaboration and information exchange. The major driver of Net-Centricity has been the shift from platform-based computing that proliferated with the spread of personal computers to network-based computing occurring with the significant growth of the Internet, intranets, and extranets.

Net-Centricity represents the interdependency of people, process, policy, and technology in information exchange. People, processes, policy, and technology all must adapt to support the new Net-Centric environment, and all four must be addressed in a fully integrated manner in order to achieve Net-Centric goals.

Mr. John P. Stenbit, former DoD Chief Information Officer, described a need for the DoD to move toward Net-Centricity and Net-Centric warfare and operations. Mr. Stenbit's vision is an environment where "people throughout the trusted, dependable, and ubiquitous network are empowered by their ability to access information and recognized for the inputs they provide." Complementary to Mr. Stenbit's vision, Dr. Margaret Myers, Principal Deputy to the Assistant Secretary of Defense, Deputy DoD Chief Information Officer, describes the basic tenets of Net-Centricity as:

- A robustly networked force improves information sharing;
- Information sharing enhances the quality of information and shared situational awareness; and,
- Shared situational awareness enables collaboration and self-synchronization, and enhances sustainability and speed of command.

By taking advantage of the decreasing costs of bandwidth; information processing; and storage, the DoD is shifting or transforming from the "smart push" to a "smart pull" paradigm. Under

this new paradigm, war fighters will have the capabilities to pull information they have determined that they need to complete their objectives. Net-Centricity will allow the war fighter to pull this information from many sources including the DoD, other government agencies, and allied powers, regardless of space, time, and location.

The Defense Information Systems Agency (DISA) is currently working to field a set of capabilities that help provide ubiquitous access to reliable, decision-quality information through a net-based web services infrastructure.

The SOA Foundation product line currently consists of six capabilities: Machine-to-Machine Messaging, Service Discovery, Identity Management, Metadata Services, Mediation, and Service Security.

2.2 Service-Oriented Architectures and the Cloud

The emergence of web technologies triggered a major paradigm shift in distributed computing, the adoption of SOA, one of the key enablers of the Cloud. With SOA, user applications deal with a set of network-accessible operations and associated resources that perform services. The Cloud takes these architectural principles of SOA and realizes them in Software as a Service (SaaS) and integrated components/services in Platform as a Service (PaaS).

Services offered are described in a standard fashion, published to a service registry, discovered by a service consumer, and then invoked (bind) by a service consumer or software agent representing a consumer. Figure 1 illustrates the three steps of the common "Publish, Find, Bind SOA" model. This high-level overview outlines the key interactions in a SOA. These concepts are explored in more detail later in this document.

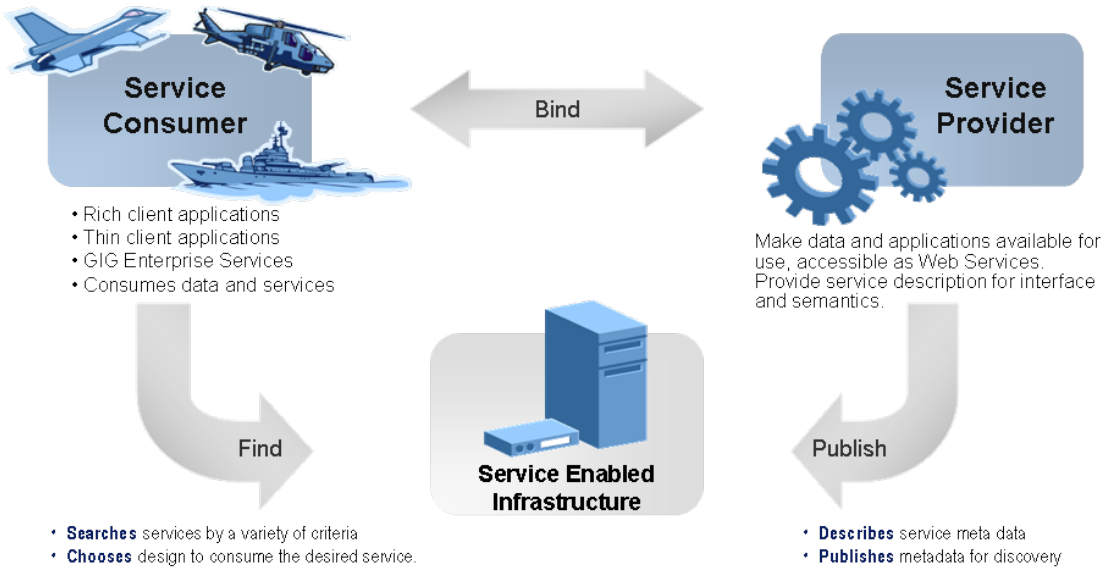

Figure 1: Service-Oriented Architecture

SOA offers several distinct advantages over traditional distributed computing technologies:

- **Maximum Interoperability:** The World Wide Web Consortium (W3C) and Organization for the Advancement of Structured Information Standards (OASIS), among others, have defined many web service standards revolving around XML based web services and HTTP-oriented REST-based web services.

- **Loose Coupling:** Web service standards define the functional interfaces that represent the minimal understanding between service consumer and service provider. Knowledge of service providers and service offers are discovered dynamically from a service registry rather than statically coded in the client program.

- **Ubiquity:** Web service protocols represent the *least common denominator* of network protocol stacks and make it easier to overcome firewall and infrastructure constraints. Web services are likely to be the most viable option for inter-agency information sharing among different autonomous networks.

The migration toward more agile SOA environments is not merely a technology push; there are also a number of key business drivers at work. There is a growing need for increased integration and collaboration across organizational boundaries. Here are some domain scenarios:

- **Supply Chain Integration:** As businesses strive to keep costs down and become more agile in meeting customer demands, it is necessary to have a technology infrastructure that can enable deep integration in the supply chain. Within this scenario, complex systems such as Customer Relationship Management (CRM) and financial systems from manufacturers, suppliers, and distributors can retrieve information and conduct business transactions with one another. For example, a business in the market for a product could shop instantly around the globe for suppliers that meet purchase requirements and dynamically negotiate deals.

- **Counter-Terrorism:** There is a pressing need in the intelligence community to provide a highly scalable system that supports collaboration, analytical reasoning, and information sharing among multiple DoD intelligence, and federal agencies. Furthermore, obtaining accurate and timely counter-terrorism intelligence requires processing unprecedented amounts of data, possibly in petabytes, from classified, unclassified, structured, and unstructured sources. There is no single system that can achieve this task and therefore the solution involves many distributed, decentralized systems.

- **Tactical War Fighting:** Similarly, in the defense sector, there is an increasing need for a Command, Control, Communications, Computers and Intelligence (C4I) system that provides a single, integrated ground picture of forces deployed to the theater. War fighters need access to real-time information and must operate within the communications infrastructure of existing global networks. Intelligent agents, for example, may automatically discover and correlate data streams relevant to a current tactical position.

It is impossible to adopt one single platform, programming language, or protocol that fulfills the needs of these scenarios. A successful architecture must accommodate heterogeneity and support interoperability in three dimensions: horizontal (across peer systems), vertical (among different organizational levels), and temporal (along a system's evolutionary path). The unique capabilities that come with service-oriented environments can successfully balance these competing dimensions.

SOA principles and supporting technologies have been the key enablers of the vision of net-centricity and cloud computing. With few exceptions, the mission functions performed by Registry Services, including the core infrastructure functions, are platform- and implementation-independent specifications that abstract underlying solutions. As a result, Commercial Off-The-Shelf (COTS) and Government Off-The-Shelf (GOTS) implementations of services can behave the same. That is, given a sufficiently robust specification, it is possible to build adaptors to current and future technologies without impacting current integrations. From the system perspective, changes in implementation matter little because they are largely invisible.

Moving toward a service-oriented, specification-driven architecture allows for the mission functions to become commodities, allowing implementations to be tailored to local environments, allowing deployments to be more or less robust based on expected load, and ensuring that vendors compete on price, reliability, and speed, not features. Net-Centricity values capabilities over implementations, and provides mechanisms that allow each member of the user community to become a catalyst of change. Services based on these principles are reliable, fault-tolerant, secure, and provide unique capabilities that enhance both the structure and substance of the network.

The DSE utilizes principles of service-orientation, and facilitates information sharing across the DoD, integrating with SOA-based services and enterprise services within PEO-ESD. DSE simplifies the publication and discovery of data services, and can be a key enabler for integration, search, discovery, interoperability, and governance of apps, services, and components critical to the DoD cloud environment of the future.

2.3 Registry Federation

The DSE has been designed to support federation with other registries across the DoD. In Federation, a user can find content contained in both a local registry and other registries that are associated through the Federation. Federation is achieved by registries (including the DSE) implementing services defined by the DoD and IC Content Discovery and Retrieval Integrated Process Team, which has defined several levels of federation. In the implementation found in the DSE, the metadata content is not replicated; the search is executed at the local registry as well as federated registries. The user can then browse the results of the search. If the user requires more information or artifacts from the federated registry, the user is required to have appropriate account privileges on that registry, and can access the metadata and artifacts directly. The DSE is expected to support at least the following federations:

- With other Core Enterprise Service metadata engines such as the Enterprise Catalog and Enterprise Search

- With structural metadata repositories and tools fielded outside PEO-ESD
- With Test and Evaluation environments put up by various communities.

Registry Federation is illustrated below.

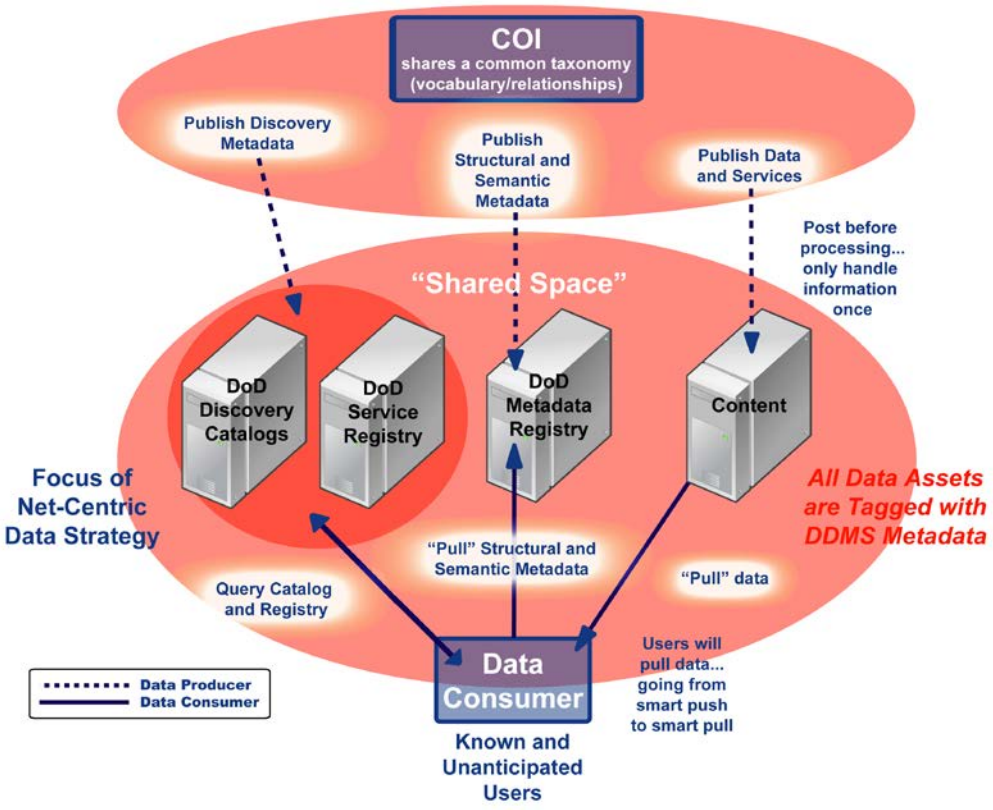

Figure 2. Registry Federation

2.4 Document Organization

This document is organized as listed below:

1. **Mission and Scope**
2. **Background**
3. **Construct**
4. **Core Functions**
5. **Summary**

3. Construct

The DSE will exist at the DoD enterprise level and is focused on the registration, sharing, and reuse of structural, semantic, and descriptive metadata about systems, services, and data resources for the DoD, US Federal Government partners, and coalition partners. It serves as a repository of this metadata, and can be linked to metadata contained in other partner repositories. It is a primary resource for the satisfaction of DoD requirements found in DoD Directive 8320.02D.

4. Core Functions

The Data Services Environment (DSE) simplifies the publication and discovery of data services and facilitates information sharing across the Department of Defense (DoD), providing an enhanced web dashboard, unifying access to many service capabilities and simplifying access for users and providing Single Sign-On (SSO) to enterprise capabilities. The DSE simplifies publication, discovery, and access to data services for information sharing, and it revolves around the goals of DoD Net-Centric Data Strategy, requiring that data be made visible, accessible, understandable, trustworthy, interoperable, and responsive to user needs on a continuing basis. For this reason, the DSE will evolve to incorporate enterprise capabilities into the DoD environment to meet these goals.

Today's data-dense environment has made the existence of the registry more important than ever. As the vocabularies have become richer, and as the number of services, widgets and apps have increased (in the form of SOAP-based services, REST-based services, and lightweight/heavyweight widgets for various platforms and mobile devices), the amount of metadata related to their use and operation has dramatically increased. The popularity of the "app store" model has underscored the need for a mechanism for registration, discovery, integration, and governance of various applications that are found in today's registry offerings. In addition to storing simple bindings and interface descriptions (which was the state of the registry a decade ago), there is now a critical need to store and make available for discovery various information, such as semantically-rich service descriptions, service interaction descriptions, expressive messaging and communication policies, and run-time service policies. The DSE aims to meet these needs.

The DSE supports the definition, publication, management, and discovery of: 1.) Structural metadata that net-centric systems use to enable discovery, share content, and otherwise interact, 2.) Descriptive metadata that describes operational and technical capabilities that can be discovered, understood, and leveraged to build warfighter and business capabilities, 3.) Semantic metadata that describes information, enhancing the understanding and interoperability of information resources within and between communities, and 4.) Metadata describing the provenance and suitability of information for use across the department.

4.1 DSE Capability Overview

A high-level overview of the function of the DSE is shown in Figure 3 – focusing on service management and metadata management.

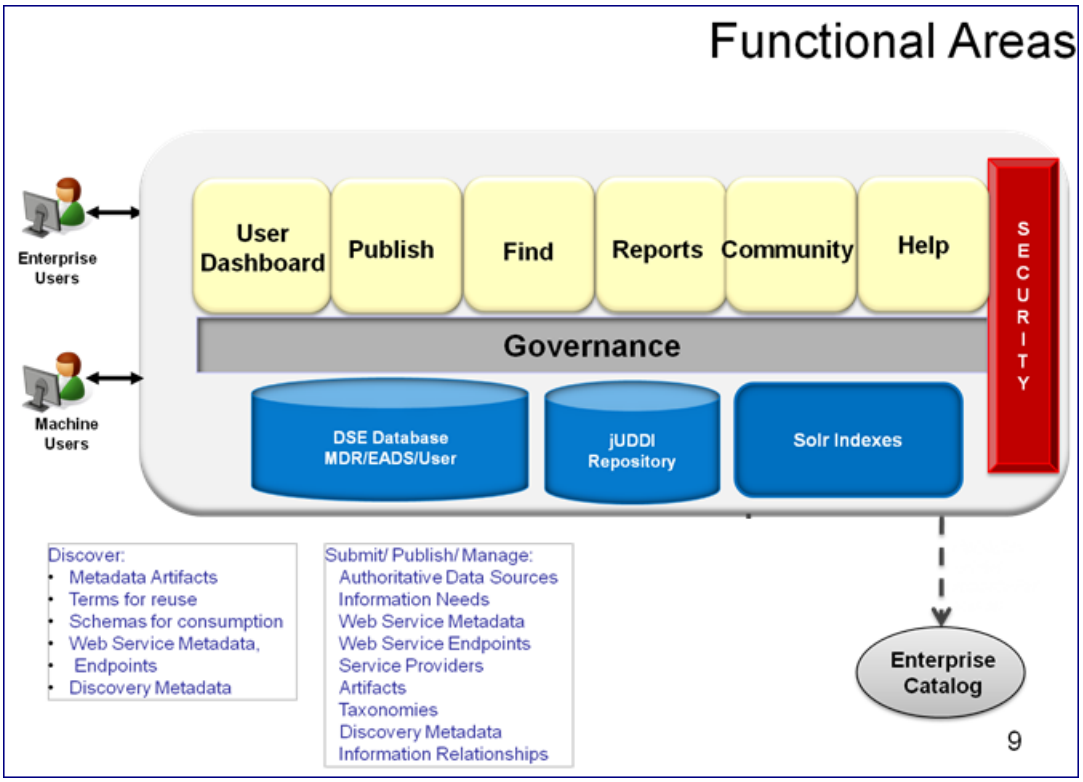

Figure 3. DSE Functional Overview

DSE currently provides services related to publication, discovery, governance, community activity, and user help:

- **User Dashboard** (Labeled Home) is the starting point for authenticated users for accessing the DSE and provides the initial user experience following authentication. It is designed to provide comprehensive information on the users' recent actions at a glance and serve as the launching point for subsequent actions. Most users will use the Home capability to serve as a navigation focus for accomplishing their subsequent actions.

- **Find** provides the capability for searching the DSE for content related to metadata about data, services, and authoritative data sources. Users can discover all types of metadata (without the need to understand the methods and organizing concepts involved) using a keyword search or using advanced search capabilities to locate metadata that meets their needs. This includes structural metadata artifacts such as XML Schemas, WSDL files, Schematron files used to validate XML files, eXtensible Access Control Markup Language (XACML) access control policy files for services, XSL stylesheets, Web Ontology Language (OWL) taxonomies, amplifying documents, sample XML documents, source code, domain value documents, and other data and service assets required to improve development efficiency through the discovery and reuse of data and service resources.

- **Publish** allows users to quickly publish metadata and related artifacts to the DSE, advertise services and other capabilities, assert data needs, advertise authoritative data

sources, and ensure compliance with Data and Service Strategy Objectives. This includes structural metadata artifacts such as XML Schemas, WSDL files, Schematron files used to validate XML files, eXtensible Access Control Markup Language (XACML) access control policy files for services, XSL stylesheets, Web Ontology Language (OWL) taxonomies, amplifying documents, sample XML documents, source code, domain value documents, and other data and service assets required to improve development efficiency through the discovery and reuse of data and service resources. Simplifying integration with lower-level services that integrate with registries and other services, it provides users with a simplified workflow for publishing metadata to the DSE and removes the complexity for the user by providing a clear set of workflows and a single interface point for publishing metadata. As part of publishing a resource the DSE can also produce a DoD Discovery Metadata Specification (DDMS) record that is pushed to the Enterprise Catalog to facilitate information resource discovery. DDMS records can be produced automatically for services as part of the service registration process; in addition the DSE includes a DDMS record composer that can be used to generates and publish a DDMS record for any desired information resource.

- **Governance** functions allow users with appropriate roles to oversee and manage the information assets they are responsible for and manage their visibility, status, and provenance from the DSE home page. A user with multiple asset management roles can manage all assets from a unified interface. Structural and Semantic metadata and artifacts, which historically were published to the Metadata Registry (MDR), are organized into administrative namespaces that have a Namespace Manager and Points of Contact that perform routine oversight of the namespace and are responsible for performing basic quality assessments of submitted artifacts. Descriptive metadata about web services and other capabilities, which historically were published to Service Discovery, are managed by service providers, who can designate others as Publish Decision Authorities or Authorized Service Providers, who must approve the descriptive metadata for it to become visible. Descriptive metadata about data needs and ADSs are managed by Authoritative Bodies, which can have managers, Points of Contact, and Administrators, who approve data needs, system designations, and authoritative data source nominations.

- **Reports** provides users with a number of pre-defined reports that can be run on demand based on user specified criteria, simplifying the effort needed to review and assess the status of metadata and artifacts in the DSE. The output of reports can be downloaded for further analysis.

- **Community** includes resources that allow users to access collaborative functions that facilitate information exchange and sharing across the user base. The Community resources include a Community of Interest Directory, introductory content explaining the DoD Net-Centric Data Strategy, Net-Centric Service Strategy, implementing policies and directives which build on the concepts introduced in the Background section of this CONOPS.

- **User Help** provides user manuals, CONOPS, tutorials, and other support information intended to facilitate self-guided training and day-to-day utilization of DSE capabilities.

The DSE also includes mechanisms for users to provide feedback and support on every page, with links to top-level help documentation as well as focused tutorial content on many capabilities.

An important capability of the DSE version 2.0 is the unification of user functions so that a user doesn't need in-depth understanding of the DSE and its technical implementation to use it effectively.

The top-level DSE menu capabilities are described in the following sections.

4.1.1 User Dashboard Overview

A screen shot of the DSE User Dashboard (Home) is shown in Figure 4. Users authenticate using approved certificates or user id and password. After authentication, the DSE provides a dashboard that provides access to the various DSE capabilities, providing a unified look and feel regardless of which capability you are using. By default, the Home screen provides immediate, at-a-glance access to Find, Publish, and Publication Management capabilities, as well as system announcements, a list of events, and a Welcome panel with selected links for common actions to assist a new user with getting started. This home capability can be customized by the user to remove or add content to provide direct access to frequently used functions from this screen, selecting the information that is displayed upon login and most relevant to them. The user can initiate most common actions directly from this home screen, such as Find and Publish without the need to navigate to the specialized page.

Beyond these common actions, the user can access all of the capabilities of the DSE from the top menu bar. These capabilities are addressed in subsequent sections.

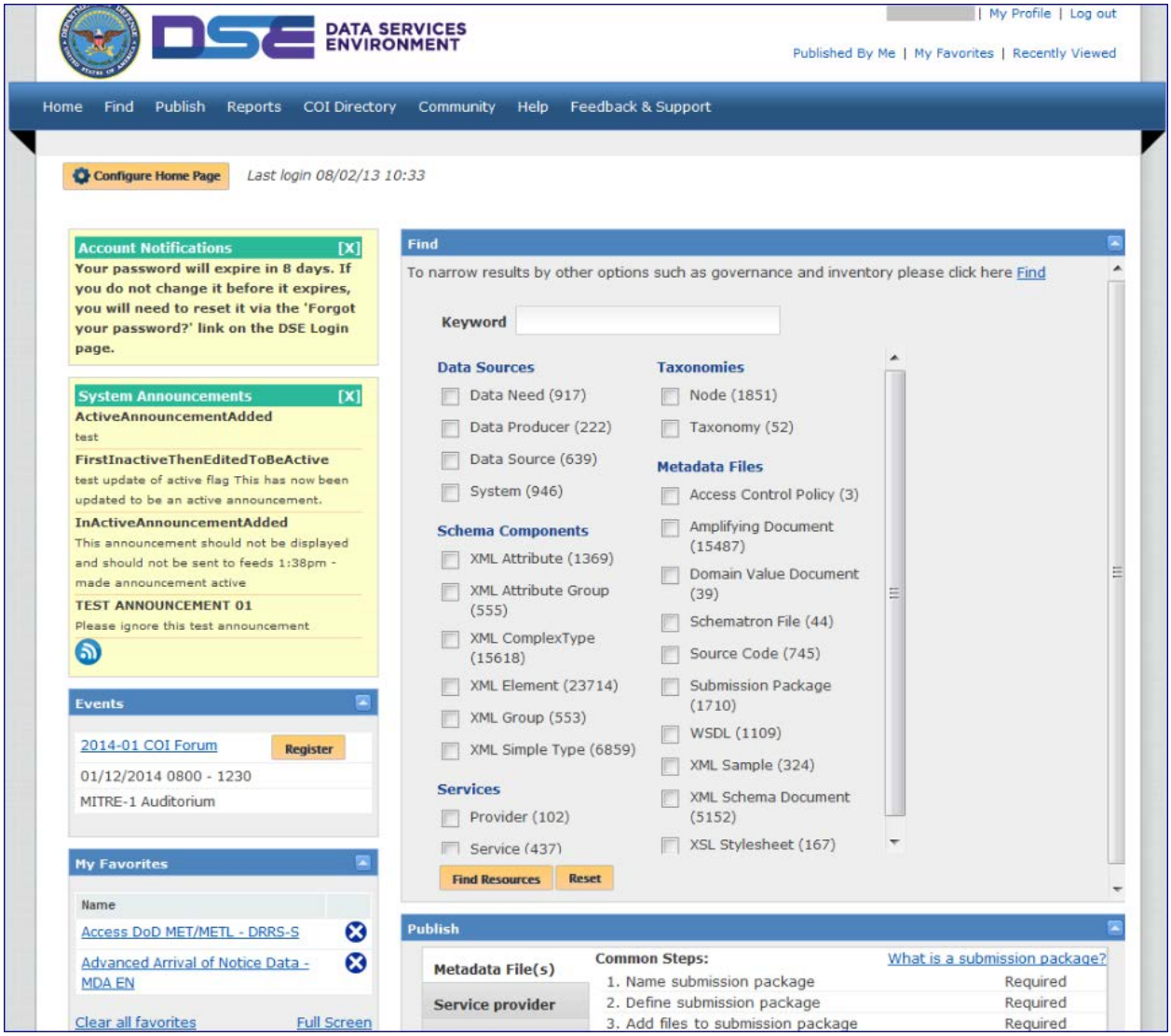

Figure 4. DSE 2.2 Home Capability

From this page the user can initiate any actions needed, from finding metadata resources to governing metadata that has been published, either directly on the dashboard or through the top menu bar with links to the available functions.

DSE currently provides unified access to metadata and metadata artifacts related to structural metadata, services, and authoritative data sources, and in the future DSE can provide further integration to mission-critical capabilities, such as Identity and Access Management (IdAM) infrastructure for providing authentication and Attribute-Based Access Control (ABAC) to services published through DSE. As we enter the era of Cloud and distributed computing, the DSE will be a driving force for integration, discovery, interoperability, and governance of apps, services, and components critical to the DoD cloud environment. The DSE is intended to make common re-usable metadata artifacts, such as WSDLs, schemas, and taxonomies, available to other data producers across the enterprise for reuse or integration/ incorporation into their solutions. This functionality could enable cost efficiencies across the enterprise application developer communities. Individual capability developers will no longer be required to define,

configure, and maintain a data environment for that capability's use. Leveraging shared common data will decrease development cost, deployment time, and security assessment effort, and will also reduce shared support costs for individual data solutions, and DSE would increase effectiveness of the capability by introducing value-added layers to the data, such as data processing, output or mediation across vocabularies.

4.1.2 Find Metadata Overview

4.1.2.1. Metadata Search and Discovery

Users are able to discover all resources submitted to the DSE, including Structural and semantic metadata, descriptive metadata about services and Authoritative Data Sources. Users are able to search for information resources through the Find page, which provides several fields that can be used to refine a query. Once resources are discovered, users may browse the resources for additional information and relationships to registered taxonomies and other metadata. Users may also be alerted of changes or modifications to those resources in which they are interested by subscribing to the resources.

4.1.2.2. Access, Retrieval and Subscription

Once an information resource is discovered in the DSE, it can be retrieved. Users can locate the package to which the information resource of interest belongs and download the package, or retrieve the specific artifact directly. Retrieval is also possible through DSE web services. For detailed information on downloading packages and leveraging the DSE's web services, see the *DoD Data Services Environment User Manual* and the DSE's Web Services materials respectively. Once downloaded, resources may be used in several different ways, e.g., employed in transaction formats, or used as a structure for caching in support of various specialized metadata engines. These generally involve ingesting all or part of a given resource into a tool or process not provided by the DSE. Users can subscribe to the resource, which results in a notification when the resource or its status changes.

4.1.2.3. Understanding Metadata

The DSE provides functionality that facilitates the understanding of metadata products. The expectation is that enhancing the understanding of metadata will support improved linkage of metadata concepts, as well as better metadata specification. Examples of how the DSE supports the understanding of metadata include:

- Finding terms and their definitions used in published schemas
- Finding concepts related to searched terms
- Easily accessing descriptive metadata for services/ capabilities, systems, organizations, service providers, authoritative bodies, data needs, and data sources.

4.1.2.4. Identifying Metadata Linkages

Information interoperability is facilitated by identifying linkages, interdependencies, and subsequently building on discovered metadata through element and term reuse. The DSE provides mechanisms for asserting and viewing the relationships between metadata artifacts and terms, which can then be leveraged in subsequent metadata artifacts. Linkages can also be

asserted through subscriptions (being informed of changes to artifacts or descriptive metadata) and statements of dependencies (data source assertions, service dependencies, comments, favorites).

4.1.3 Publish Metadata Overview

The DSE contains the structural and semantic metadata artifacts critical to successful development, operation, and maintenance of existing and future capabilities that support the DoD Net-Centric Data Strategy. The DSE embraces the vitally important features of versioning, access control, lifecycle management, and portal publishing. It is the Department's primary tool for enabling transparent, globally distributed, and collaborative management of structural metadata. The DSE also supports the establishment of relationships between metadata artifacts for mediation and convergence on common representations, a feature that will prove progressively necessary in the future. In addition to human users, the DSE's growing capability to support machine-to-machine transactions will provide ways of rapidly mediating, validating, and synchronizing widely distributed metadata online in near-real time.

The DSE is designed to interact with other Core Enterprise Services including People, Service, and Content Discovery, as well as the Mediation and Security services. It is also critical to the successful deployment and interoperation of specialized COI capabilities supporting service-oriented architecture, data management, various application packages, and other key information sharing activities that depend upon a set of reliable metadata resource management capabilities. The DSE is a general purpose hub that supplies structural metadata to numerous specialized metadata capabilities. It is intended to be the authoritative source for DoD metadata.

It is a virtual “place” where collections of metadata components, in which DoD organizations and others have invested, can be:

- Published, made visible, accessible, and understandable to large National Security audiences
- Transparently and collaboratively evolved and otherwise managed by representatives of a large, diverse, geographically distributed group of people and organizations
- Monitored to determine contextual relevance (importance/priority), quality, usage and other factors that affect engineering and resourcing (future investment) decisions
- Exploited by machine-to-machine process in support of such functions as validation and transformation.

The DSE serves all of DoD and some users outside the Department. It reaches everyone regardless of the security enclave in which a customer (machine or human) may be located. There are two fundamental user types: (1) Publishers who post and configuration manage metadata artifacts in the DSE so those artifacts can be reused across the Department to support information sharing; and (2) Consumers who retrieve and subscribe to metadata artifacts or information about artifacts in order to use the artifacts to create information sharing capabilities. Other “hands-on” users include DSE operators/administrators, program planners, testers, modeling and simulation engineers, and several echelons of acquisition overseers.

4.1.3.1. Publication (Submission)

Users are able to publish Information Resources of various types to the DSE, such as XML Schemas, WSDL files, Schematron files used to validate XML files, eXtensible Access Control Markup Language (XACML) access control policy files for services, XSL stylesheets, Web Ontology Language (OWL) taxonomies, amplifying documents, sample XML documents, source code, domain value documents, and other data and service assets required to improve development efficiency through the discovery and reuse of data and service resources. In addition to these artifacts, the DSE also captures descriptive metadata about services and Authoritative Data Sources that can reference those artifacts.

The DSE uses wizards to guide the user through the metadata submission and publication process to reduce the complexity of metadata publication. These wizards also ensure that metadata is as consistent and complete as possible, while minimizing user re-entry of data.

4.1.3.2. Schema Validation

The DSE has provided Schema file validation (validating form and import integrity) for some time. It now also provides a web-based capability for validating XML documents against schemas in the MDR with Schematron files (either in the registry or uploaded).

4.1.4 Reports Overview

DSE includes a powerful web-based report-generation tool that creates reports of various types (artifacts with the most similar relationships, most downloaded, most subscribed, recently approved artifacts, and many others) based on namespaces and resource types. When a report is selected, the user fills in the needed parameters and the formatted report (as a PDF file) is delivered to the user. Additionally, ad-hoc reports can be generated using the Find capability, which generates a table result with user-selected content that can be downloaded in either an XML or Comma Separated Value (CSV) format for viewing or analysis in a spreadsheet.

4.1.5 Communities of Interest Overview

Communities of Interest (COIs) is a term used to describe any 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 organized and managed outside of the DSE; the DSE provides mechanisms for both discovering COIs and providing COI-originated metadata either as information resources (such as a COI taxonomy or descriptive information resource metadata).

Users can add or update COIs, view domains and mission areas, and view COI links. COIs can indicate an interest in administrative namespaces, permitting users to better understand the operational and management structures that govern metadata, while serving as an organizational structure for management of taxonomies, services, and more generally, any information resource. COI definitions can include descriptive metadata, including Domains and Mission Areas.

Domains are defined as subsets of mission areas and represent a common collection of related, or highly associated, information capabilities and services.

Mission Areas are a defined area of responsibility whose functions and processes contribute to accomplishment of the mission. Generally, COIs are formally defined as chartered by or supporting one or more domains.

4.1.6 Community Overview

The DSE supports a number of capabilities intended to facilitate information exchange with the user community and between members of that community. These are accessible from the Community page of the DSE. They include the announcement of events of potential community interest, such as Metadata Working Group (MWG) meetings and visibility of archived briefings and reference documents, and links to user forums hosted on forge.mil. As the DSE evolves and user input is received other content will be added.

4.1.7 Help Overview

The Help capability provides access to user guides, tutorials, frequently asked questions, and links to related content of interest to the metadata community. The intention is to provide material that guides both first-time users who need detailed information on getting started using the DSE as well as power users who need to use less common but powerful capabilities. The intent is to expand this content to incorporate use questions and feedback.

4.1.8 Feedback and Support

The Feedback and Support capability facilitates users contacting the DSE administrative team with their questions, observations, and recommendations for improvements. Information on the user submitting the feedback is automatically populated.

4.2 Descriptive Metadata about Authoritative Data Sources

An Authoritative Data Source (ADS) is a recognized or official data production source with a designated mission statement or source/product to publish reliable and accurate data for subsequent use by data consumers. An Authoritative Data Source may be the functional combination of multiple, separate backend databases, data feeds, systems, or other data sources. The diagram below illustrates the core concepts making up an Authoritative Data Source.

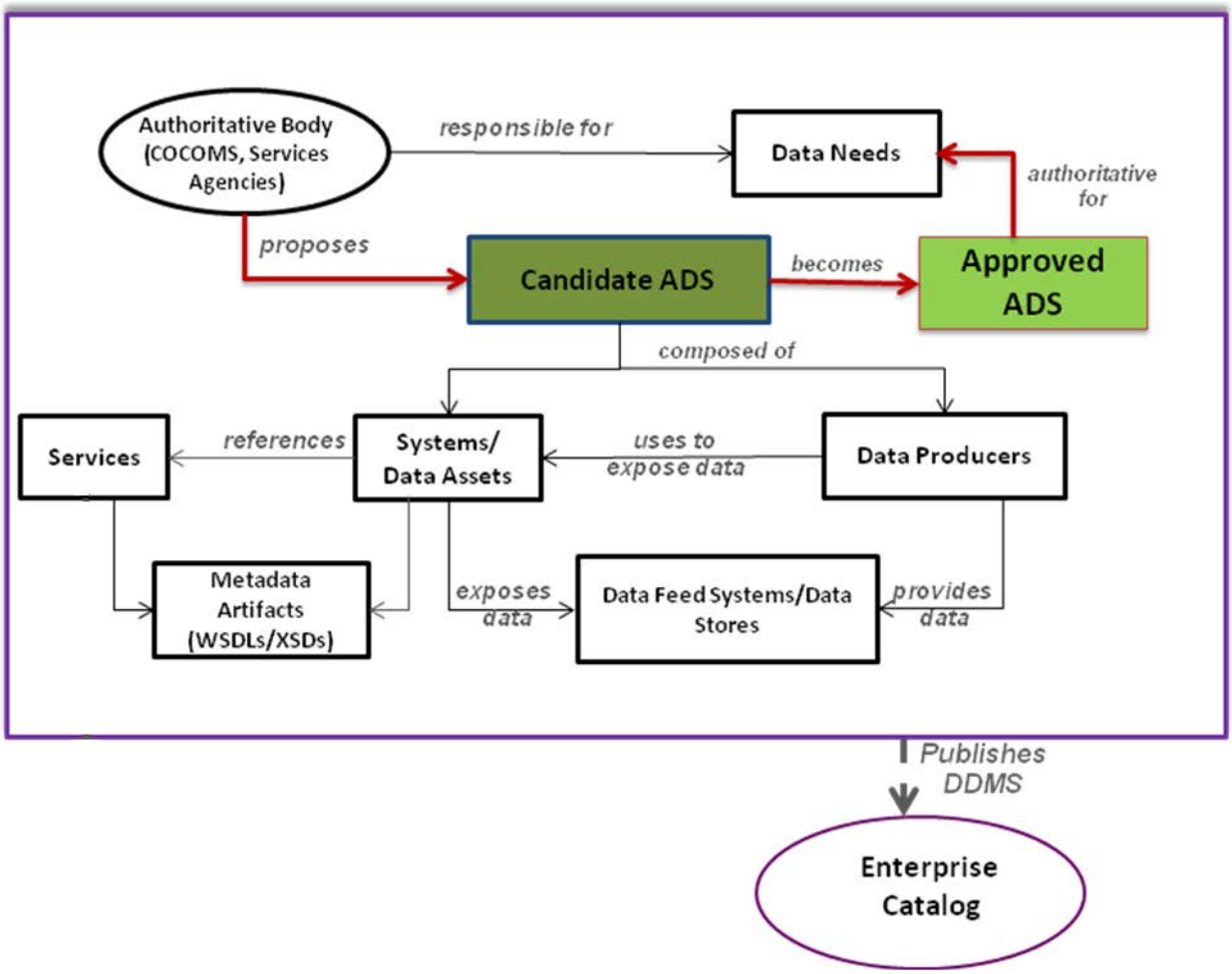

Figure 5. ADS - Core Concepts

As illustrated in the figure above an authoritative body is an entity responsible for 1) meeting or fulfilling one or more specified data needs; and 2) identifying candidate data sources that might be designated as authoritative for one or more data needs for which the authoritative body is responsible. The DSE does not set the criteria for being deemed an authoritative body, but merely allows authoritative bodies to identify the data needs for which they are responsible and identify authoritative data sources that provision data fulfilling those needs.

A data need is a named and defined specification for a particular type of data that supports one or more operational requirements. It may be generic or very specific, however must be able to be satisfied by a single data source.[1] It is important to note that the concept of a data source is not intended to refer to a single system, or family of systems, but the combination of data producers, i.e. the operational providers of the data needed, the systems or tools that those data producers use to expose that data, and any underlying databases, data feeds, or systems that are also exposed through those systems or tools.

[1] A single data source, however, may be made up of numerous other systems, services, databases, data feeds, or capabilities.

The specified combination of a system or tool, a data producer, and a data need form the basis of a proposed (or candidate) Authoritative Data Source. The DSE does not make assumptions about the combination of these entities but supports the binding of them by operational users. The act of binding, i.e. specifying a combination of these concepts, results in the creation of a candidate or proposed Authoritative Data Source which is then brought to the attention of all authoritative bodies that have specified the bound data need as within their area of responsibility. The authoritative bodies or their designees then may approve the proposal thus formally creating an Authoritative Data Source.[2] The diagram below further illustrates the process surrounding the establishment of an Authoritative Data Source in the DSE.

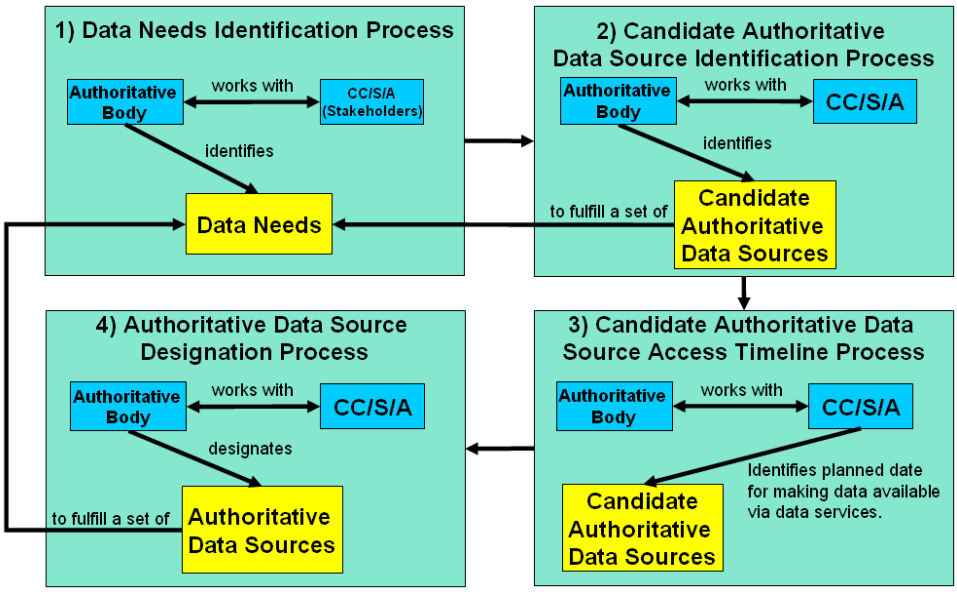

Figure 6. Process of Establishing an Authoritative Data Source

4.3 Descriptive Metadata about Services and Capabilities

In order to achieve Net-Centric Data Strategy goals, the nature of each asset must be described, its location identified, and its structure/significance made known. This metadata is the information about an asset that provides an asset's meaning and technical characteristics. The creation, publication, exploitation, and management of this metadata form the lynchpins of the DoD's Net-Centric Data Strategy – indeed, net-centric interoperability across the DoD hinges on metadata exposure, exploitation, and reuse.

As the DoD moves into the Cloud, it becomes increasingly important that metadata be visible, discoverable, and managed. The nature of service-orientation, one of the key enablers of the Cloud, requires rich metadata environments to support service and data asset lifecycles from concept, design and implementation to operations to retirement. The DSE provides the core

[2] In the event that multiple authoritative bodies are responsible for the same data need(s) the DSE allows for those authoritative bodies to act autonomously when approving proposed Authoritative Data Sources. That is, if authoritative body A chooses to approve the proposed ADS and authoritative body B chooses not to, then the DSE will reflect that the data source in question is authoritative for the data need as supported by Authoritative Body A, but not for the need as supported by Authoritative Body B.

components of a net-centric metadata environment to fully and effectively enable net-centric operations for the Department. The nature and complexity of a metadata environment sufficient to support net-centric operations, has a tendency to be daunting to program managers, architects, and developers attempting to understand the intricate relationships between the components of the metadata environment. There is now a critical need to store and make available for discovery various information, such as semantically-rich service descriptions, service interaction descriptions, expressive messaging and communication policies, and run-time service policies.

The DSE provides easy-to-use mechanisms for making service and capability metadata visible, accessible, and understandable, as required by the DoD's Net-Centric Service Strategy. These capabilities streamline the process of publication and discovery, executing the extremely complex low-level service interactions with various enterprise components, providing an easy-to-use user interface abstraction to the user.

These service discovery and publication capabilities of DSE play a critical infrastructure role:

- They allow service providers to publish/advertise service specifications, metadata, and accessibility. Service providers in the DoD may offer web services, cloud services, or applications (such as service-enabled target tracking applications), data repositories (such as a Coalition Shared Database), devices (such as sensor platforms), or even non-technical business functions (such as a Helpdesk for technical support).

- They allow service consumers to discover information as advertised by providers. The consumers may include humans and service consumer agents such as thick clients (for example, service-enabled Command and Control applications), thin clients (for example, Web browsers), or devices (for example, smart phones).

- They allow end-users to discover services or service providers by performing a web-based keyword search or using the advanced search capabilities to locate the appropriate services/providers to meet their needs. In addition, users may add, edit, and get notified on subscriptions they have to providers and services.

- They allow enterprise service developers to transparently enhance discovery, retrieval, and publishing services without interrupting normal business operations.

They provide governance over user-defined taxonomies and fingerprints, and provide enhanced oversight and management of publications.

4.3.1 Benefits of DSE's Publication and Discovery Capabilities

The benefits of the publication and discovery capabilities are numerous:

- **Promotes reuse:** If developers can locate information in a service registry such as service offers, service specifications, and taxonomical information, they can readily reuse these existing entities to save time and avoid duplication of effort.

- **Enables governance:** By utilizing a common DSE capability for Service Discovery, governance models can be created and enforced to provide: quality of information,

establishment of trust between providers and consumers, predictability of service performance and discoverability, interoperability, reuse and control.

- **Enables lifecycle management:** Services may go through several stages during their life span. Entities in the service registry may mature through varying stages such as in development, operational, and deprecation. By controlling the service publishing process, the DSE helps to enforce such lifecycles.

- **Promotes unified understanding of metadata:** Because metadata may be accessible from multiple service registries, a unified understanding of such metadata is promoted throughout the enterprise, which in the long run helps promote semantic interoperability across organizational boundaries.

- **Enables collaborative development:** Service providers can create service specifications and service metadata to submit to service registries for use and potential enhancement by others. For example, a service provider can publish a service offer and related service specifications in the DSE; another provider may view the service specifications published with the service offer and publish another service access point that conforms to the same interface specification but with updated/enhanced functionality. This new service can be made available for discovery and use by the community.

- **Promotes service virtualization:** With the ability to discover service access point reference information (such as network address and protocol) at run-time, this information need not be "hardwired" into service consumer agents. Location transparency enables true loose coupling among components, fulfilling one of the goals for Net-Centricity.

- **Promotes service composition and orchestration:** Dynamic discovery also is a prerequisite to service composition, which is the ability to assemble multiple existing web services to form a business process or achieve a high-level business function. For example, to support a new tactical planning mission, a workflow management service discovers three available services in the network: a sensor service providing Unmanned Aerial Vehicle (UAV) imaging data, a weather service providing weather data for the target area, and a Blue Force Tracking service providing the force information. The workflow service invokes those services dynamically and processes the results to be used by the planning mission.

- **Enables dynamic run-time decision support in the Cloud.** If applications and services can locate information in a registry, they can make automated decisions at run-time. For example, an app store can group apps and services based on taxonomic classification, a web client of a Cloud service can adapt to security messaging and access control policies defined in security metadata, and federated search services can smart search categorizations based on a user's preferences and specific taxonomic classifications in the registry.

There are countless scenarios that show the benefits of Service Publication and Discovery. One scenario involves dynamic discovery of services. Using the DSE, a service in the DoD Cloud that returns air track data may dynamically discover and subsequently invoke another service

implementation that represents a sensor system that provides sensor data to the air track data service. Figure 7 illustrates this scenario in which a soldier, using an app on his handheld, utilizes a service that dynamically discovers and consumes another service that had been published in the registry.

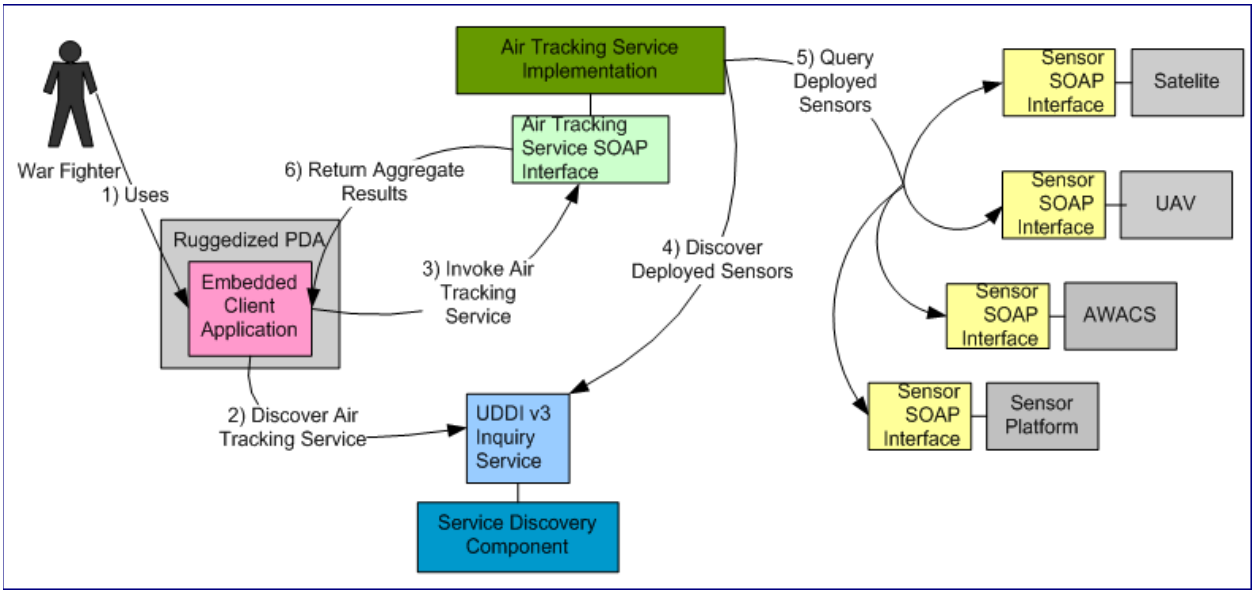

Figure 7. Dynamic Service Discovery and Composition

Figure 7 depicts the following steps:

1. A user uses a mobile app on his Handheld to query an Air Tracking Service.
2. The mobile app makes a service request on behalf of the user to the DSE's Inquiry Service. The application queries for an access point of the Air Tracking Service.
3. The application selects an access point and performs the requested query against the Air Tracking Service interface.
4. The Air Tracking Service processes the user's request and determines that it needs to aggregate data from currently deployed sensor systems. The Air Tracking Service queries the service for deployed sensors that provide the Sensor SOAP Interface.
5. The Air Tracking Service processes the results from step 4 and performs a query against each sensor system access point returned by the Service Discovery component.
6. The Air Tracking Service implementation processes the results from the queried sensor systems and returns the results to the embedded application. The application processes the results and displays the data to the user.

This is but one example of the benefits of publication and discovery in the DSE.

4.3.2 Service Publication

Discovery of metadata about services is dependent on the *publication* of the metadata. Service information entities that are published to the registry undergo a series of publishing stages as

their status and usage progress. The defined publishing stages for Services and Capabilities are illustrated in Figure 8.

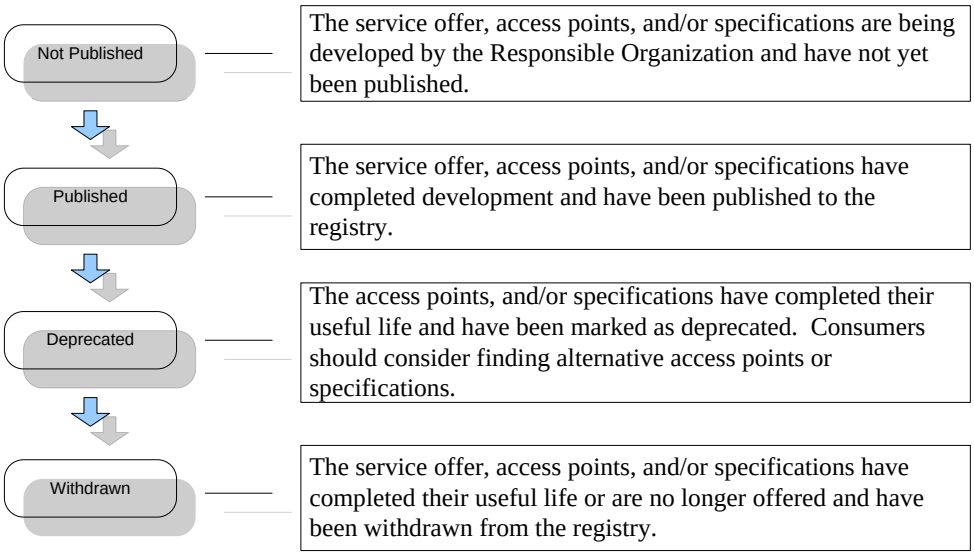

Figure 8. Stages of Service Information Entity Lifecycle

Service publishing involves placing service information entities such as service providers, services offers, service access points, service specifications, and the technical "fingerprints" of the each of these entities into a registry. The following scenarios need to be supported:

- **Static (Design-Time) Service Publishing**: A human user / operator serves as the publisher, who uses a Web based graphical user interface (GUI) to publish the information entities in the registry. This is the most straightforward approach and is likely to be the primary way of publishing services in the near term.

- **Dynamic (Run-Time) Service Publishing**: In this case, the publisher is an authorized application (possibly the service itself), which uses a publishing Web service / API provided by the registry to publish the service information entities.

- **Static or Dynamic Updates to Information Entities**: In addition to automated publishing, an application may need to update its definitions and metadata in the registry, so that the entities in the registry are kept in sync with the operating conditions of the real application. For example, when a service provider agent is moved to a new endpoint, or a new authentication requirement is initiated, the access point entry in the registry needs to be updated accordingly.

Some of the details of the publishing functionality are below:

- **Publishing a Service -** The complexity of publishing assets in the DSE revolves around relationships between metadata assets and various enterprise components. The DSE provides a clearly defined interface that simplifies the user experience and streamlines this process by providing a simple user interface that prompts the user for the required information in the correct order and pushes that information to the proper components of the DSE and Enterprise Catalog. In the steps above, the DSE will prompt the user for a

few required pieces of information, collect the XML Schema files, collect the WSDL file, and push those assets to the DSE if they have not already been published. It will then provide the user with the URLs for those assets and point the user to the service on-boarding process. Optionally, the user may publish a DDMS record about the service to the Enterprise Catalog using the "Advertise Resource capability."

- **Publishing Resource Advertisements** – The DSE provides a streamlined interface for creating discovery information about an asset and publishing it to the Enterprise Catalog so that it can be discovered. It provides a capability to save that information for future publication or to publish that information to the Enterprise Catalog immediately. Alternatively, the DSE provides a capability that allows users to upload DDMS 4.0.1 compliant DDMS records (i.e. XML files that validate against the DDMS 1.4.1 XML Schema) and publishes the supported discovery information to the Enterprise Catalog.

- **Publishing Metadata -** Publishing metadata is the broadest of the publication capabilities that the DSE provides. The publish metadata capability allows the user to publish just about any type of metadata asset that might exist. The publish metadata interface provides users with a simplified flow for publishing XML Schemas, XML Sample Documents, briefings, documents, source code, or any other artifact that users might otherwise publish to the DSE. In addition, the DSE collects information regarding the Systems or Programs of Record using the metadata artifact being published to aid in various DoD requirements to annotate metadata assets with the systems and programs using those assets.

4.3.3 Service and Capability Information Model Concepts

The basic information model involved in service discovery includes concepts that represent service providers and consumers, their relationships, and their roles within a service-oriented environment – to include SOA-based web services and Software as a Service (SaaS) in the Cloud.

- A ***mission entity*** is operationally responsible for performing a mission function, or operationally in need of having a mission function performed. Mission entities may include entities such as individuals (humans), organizational units, COIs, and programs of record.

- A ***service*** is the performance of a mission function. It is important to note that a service is an act and not an object.

- A ***service provider*** is approved by a mission entity as having the responsibility, authority, and financial resources to create and maintain services in support of a mission function. Initially, the Registry Services program will approve service providers on behalf of a mission entity.

- A ***service consumer*** is a mission entity that leverages a service which was created by a service provider in support of a mission function.

- A ***service offer*** is a proposal to perform a mission function for others (service consumers). A service offer is made by a specific service provider and is described by a

referenced (or embedded) service specification. In the context of the Service Registry, a service offer is the equivalent of a web service.

- A ***service specification*** is a formal description of a mission function to be offered. For a service specification to be useful, it should contain not only a textual description of the mission function, but also a precise description of the behavior of the mission function in terms of its inputs, outputs, and associated semantics. One key benefit of maintaining a service specification as a distinct data object is that it allows an enterprise to exert governance over the described behavior of mission functionality independent of the mission entity that proposes to perform the function in a service offer. In this way, when more than one provider offers to perform the same mission function, a consumer can know that two offers are equivalent by determining whether both offers refer to the same service specification. The consumer may then choose from which provider it wishes to consume the service based on other available information such as reliability of the providers, convenience of access points, or past performance characteristics. In the context of the Service Registry, the service specification describes a web service, and is registered in the service registry.

- A ***service access point*** is a logical location where a mission function is made accessible (also referred to as *endpoint*). A mission function may be offered from one or more service access points.

- A ***service provider agent*** is a combination of hardware and software that a service provider deploys to perform a mission function. A service provider agent is necessary to receive requests from, and send responses to, a service consumer agent.

- A ***service consumer agent*** a combination of hardware and software that a service consumer uses when requesting that a mission function be performed by a service provider. A service consumer agent is necessary to send requests to, and receive responses from, a service provider agent.

- A ***service provider host*** is the logical system on a network from where a service provider agent performs its function. A service provider may offer the same service from multiple hosts. Also, multiple service providers may offer different service from the same host (for example, in a managed service environment).

- A ***service consumer host*** is the logical system on a network from where a service consumer agent performs its function.

Figure 9 illustrates the relationships between the concepts in the information model.

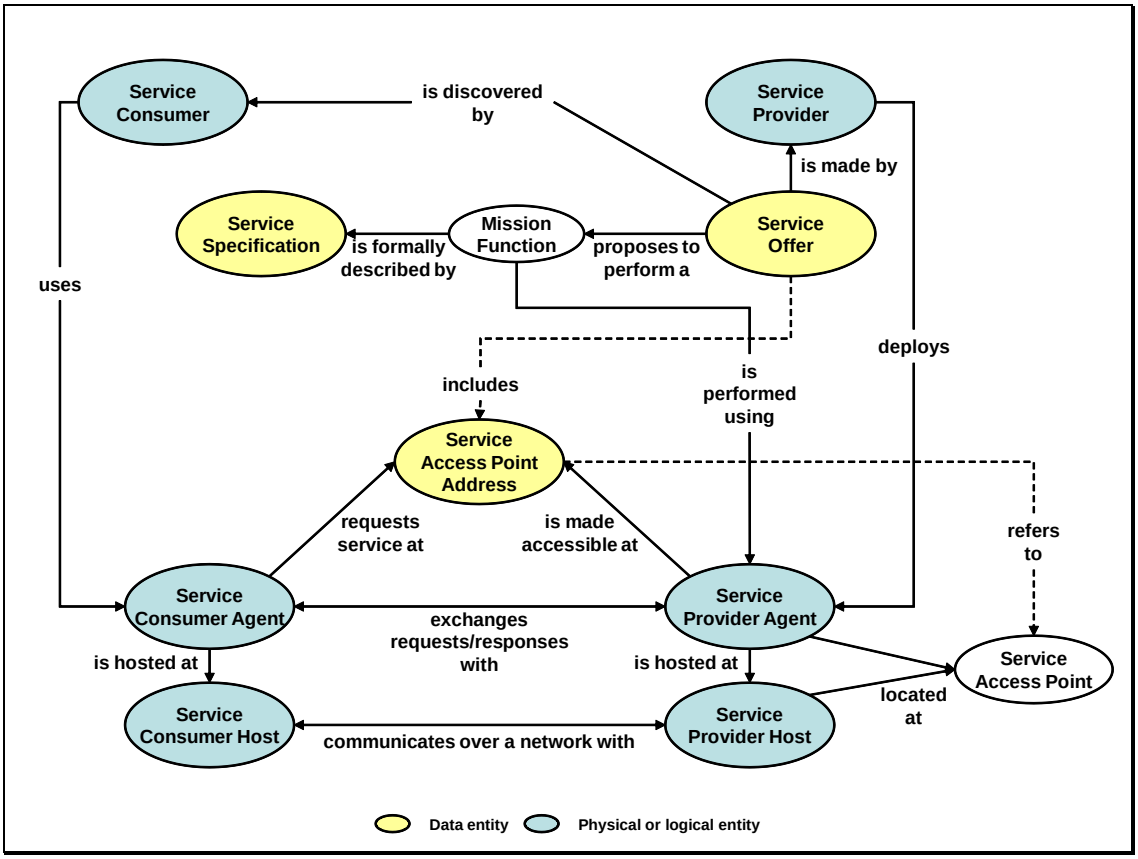

Figure 9. Information Model Relationships

4.3.4 Concept Realization

To realize the information model concepts, the DSE utilizes web services as an implementation specific approach to implementing a Service-Oriented Architecture. As part of the information model, the technical behavior of a service specification (typically in a machine process-able format) is required. For web services, this part of the service specification is generally written in Web Service Description Language (WSDL) format [WSDL-1.1]. WSDL is a W3C standard XML language to describe web services. A WSDL document describes the technical interfaces and operations available (what the web service can do), what protocol and transport should be used to consume the web service (how to consume), and the network address of the service access point (where to find it). WSDL also provides extensibility points for service providers to describe other technical and semantic aspects of their web services. For example, Web Services Policy Attachment [WS-POLICY] is such an approach that embeds security policy elements as WSDL extensions.

The ultimate goal of the DSE is to make all relevant information about web services and other capabilities available for discovery. Due to the acceptance of the WSDL standard, support for publishing and searching WSDL definition information in the Service Discovery capability becomes a requirement. As technologies evolve and change as RESTful services are deployed that have no concept of WSDL, DSE will evolve with the technology.

The use of WSDL standards can be seen in the following scenario: A technical committee defines an enterprise content search specification with an abstract WSDL document. The abstract WSDL uses a targetNamespace that reflects the technical committee that defines the specification. Organization A decides to offer its document repository to its business partners using the search specification. Organization A hires an IT implementation company B to service-enable the document repository, so it has a SOAP interface that conforms to the enterprise content search specification. When organization A offers its content search service, it prepares an instance WSDL document for its running instance. The instance WSDL document uses a targetNamespace that reflects organization A. The instance WSDL document also references the port type and bindings defined in the abstract WSDL document that comes with the content search service specification. Organization A publishes its content search service instance in the DSE using the instance WSDL document. Figure 10 illustrates this scenario.

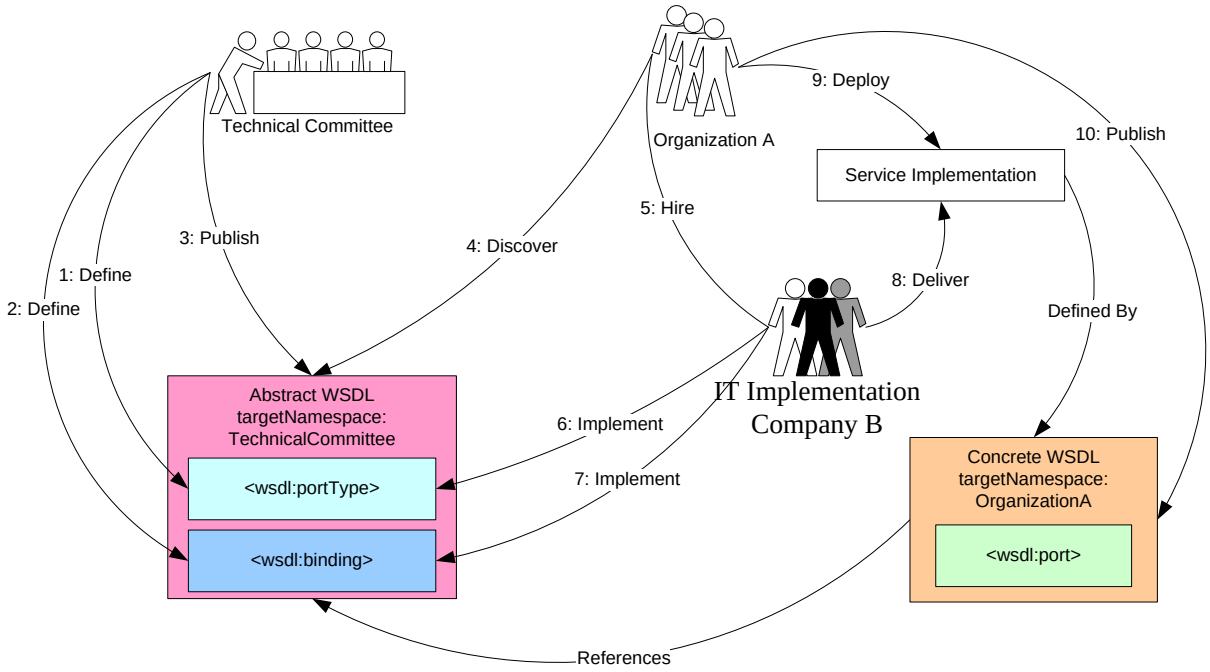

Figure 10. Abstract and Concrete WSDL Usage Scenario

4.3.5 Categorization

Information entities in the registry are categorized in a systematic manner using taxonomies. A taxonomy is a hierarchical arrangement of categories or classifications. Taxonomies are highly important for registries at large, as they provide a means to organize registry contents for efficient maintenance and discovery. For example, a user or process may wish to discover all service offers that perform data conversion operations. This query may be performed by accessing a taxonomy that categorizes service offers according to their data operations and searching for all service offers that are categorized with a value of “Format Conversion”.
Information entities are categorized using the following types of taxonomies:

- Mission entities and service providers are categorized using organizational taxonomies. For example, the DoD agencies may constitute such a taxonomy.

- Service offers are categorized using business or functional taxonomies. For example, a data service may be categorized by the map objects it provides, such as air tracks and ground tracks. The Service Registration and Governance (SR&G) Working Group (WG) has defined a taxonomy to categorize all web services.

- Service access points may be further categorized by technical taxonomies. For example, the hosting location and Quality of Service (QoS) parameters for a particular access point may be defined in this way.

Service specifications may also be categorized by technical taxonomies. For example, the source format of a service specification can be categorized to allow consumers to discover machine and human readable specifications. A service specification categorized with a WSDL source format may be processed by machine, while a service specification categorized with a text source format is better suited for human consumption.

4.3.5.1. Modeling Service Taxonomies

As reflected in this information model, service taxonomies are not modeled as a single, monolithic tree of categories, but rather as multiple smaller, relatively *flat* taxonomies. This approach brings many benefits:

- Flat taxonomies are easier to define and manage.

- Flatter taxonomies are easier to traverse and look up, hence more user-friendly.

- Because an information entity can often be categorized along more than one dimension, having multiple taxonomies is more natural and intuitive.

New taxonomies can be added without having to modify existing ones.

4.3.5.2. Predefined Enterprise Taxonomies

To facilitate enterprise service discovery, a set of core service taxonomies has been defined for services. These taxonomies are defined by the SR&G WG. These taxonomies are not set in stone and may evolve to reflect feedback from the DoD and IC communities.

When publishing an entity in the Service Discovery component, the appropriate categorization MUST be assigned to the entity and all mandatory taxonomies MUST be included in the entity categorization.

In addition to these enterprise taxonomies, additional taxonomies may be defined by COIs and local organizations to meet their discovery needs.

4.3.5.3. Categorization of Taxonomies

The approach of defining multiple smaller taxonomies has created a management problem. When the number of taxonomies becomes non-trivial, it becomes necessary to organize the taxonomies themselves so that they can be easily discovered and retrieved. The following taxonomy of taxonomies presents a set of attributes that can be assigned to all taxonomies.

- **Requirement** - The categories that denote whether categorization under a particular taxonomy is mandatory, recommended, or optional.
  - Mandatory: Denotes a taxonomy that is a required taxonomy. If a taxonomy is mandatory, a discovery entity must be categorized using the metadata defined in the taxonomy.
  - Recommended: Denotes a taxonomy that is recommended, but not mandatory.
  - Optional: Denotes a taxonomy that is optional. A discovery entity publisher may choose to use categories from an optional taxonomy to describe entities being published.
- **Dimension** - The categories that denote the dimension of the taxonomy. A taxonomy is usually defined to address one dimension of attributes of a discovery entity.
  - BusinessOrganization: Denotes a taxonomy that categorizes by business organization structures.
  - Functional: Denotes a taxonomy that categorizes by functional capabilities.
  - Technical: Denotes a taxonomy that categorizes by technical aspects.
  - ContentData: Denotes a taxonomy that categorizes by the content or data a discovery entity is dealing with. For example, a taxonomy about the types of data a service instance is providing or manipulating falls into this category. This dimension can be treated as a special case of the "Functional" category.
- **Applicability**: The categories that denote the type of discovery entity to which a taxonomy applies.
  - ServiceProvider: Denotes a taxonomy that can be used to categorize service providers (Business Entities in UDDI), for example, a business organization taxonomy.
  - Service: Denotes a taxonomy that can be used to categorize service offers (Business Services in UDDI).
  - ServiceInstance: Denotes a taxonomy that can be used to categorize service access points (Binding Templates in UDDI), for example, a taxonomy about hosting locations is applicable to service access points.
  - ServiceSpecification: Denotes a taxonomy that can be used to categorize service specifications (tModels in UDDI), for example, a taxonomy about the types of specification document such as WSDL or Interface Definition Language (IDL).
- **Scope**: Categories that denote the usage scope of a taxonomy.
  - Enterprise: Denotes a taxonomy that is defined at the enterprise level.
  - COI: Denotes a taxonomy that is defined at the COI level.

- Local: Denotes a taxonomy that is only used within the local trust domain.

4.3.6 Versioning

The guidance for versioning of web services is defined in the Joint IC/DoD Enterprise Service Versioning Guidelines. It was produced on 11 July, 2008, and is currently submitted for approval by the Enterprise Services Engineering Review Board (ES ERB).

4.3.7 Lifecycle Management

The guidance for lifecycle management of web services is defined in a Joint IC/DoD Lifecycle Management document. This document outlines the web service lifecycle taxonomy. Each version of the service information entities may also go through a development lifecycle. Here we are more concerned with the lifecycles of service access points in the registry. Like versioning, development lifecycles may vary from service provider to service provider depending on the nature of the implementation and the methodologies and processes that are in place. Below is a list of lifecycle stages as defined by the SR&G and implemented in the PEO-ESD Service Registry.

- **Concept:** This stage indicates that the potential need for the service has been identified.
- **Developmental:** This stage indicates that the service is in the process of being built or acquired.
- **Pilot:** This stage focuses on operational testing and the evaluation of the service in end-to-end mission threads. A service in this stage will have the ability to access real-world data to support operational evaluation. This stage certifies the operational utility of a service.
- **Operational:** This stage indicates that the service is currently operational and is being maintained by the service provider.
- **Deprecated:** This stage indicates that the service is scheduled to be retired. The provider should identify a disposal date, at which time this service will no longer be available. In most cases, the service provider should also provide a migration plan for the consumers of the service.
- **Retired:** This stage indicates that the service is no longer operational and should not be used.

4.4 Registry Administration

The Registry Administrator is responsible for posting all submissions into the DSE. Posting resources allows a resource submission package to be published to the Registry under the specified Governance Namespace. Generally, posting is a threshold activity intended to insert a precautionary measure into the process so that every submission to the DSE is checked to ensure that it contains conformant resources devoid of malicious code or content. After the resource is posted, the metadata may be discovered and retrieved. The Governance Namespace Manager or

Administrator of the Governance Namespace into which the submission was posted will be able to tag the submitted resources with a status of developmental, operational, deprecated, or retired.

Like Governance Namespace Managers, Registry Administrators also have additional responsibilities within the context of the DSE. Registry Administrators make up the support staff that ensures that the DSE is available to end users 24 hours a day, every day of the year. On a monthly basis, they are also responsible for ensuring that resources are accurately replicated among the registries that exist on at least two separate networks: NIPRNet and SIPRNet. In addition to these system support functions, Registry Administrators have numerous other responsibilities that facilitate continuity of the Department's metadata initiatives. These responsibilities include:

- Creating and registering Governance Namespaces and sub-namespaces requested and sponsored by appropriate authorities.
- Managing the DoD Enterprise namespace, which is used to publish and manage information resources employed Department-wide.
- Managing the "To Be Determined" namespace that handles resources that do not appropriately fit into an existing Governance Namespace and where the registrant is not prepared to establish a new namespace.
- Maintaining, monitoring, and responding to a feedback mechanism for user questions, complaints, and suggestions related to the DSE.
- Acting as a technical advisor to the DoD Metadata Working Group and its focus groups.
- Managing the acquisition, testing, and delivery of DSE capability enhancements.
- Facilitating the metadata registration process to ensure completeness, uniqueness, and technical validity of information resource submission.

The DSE provides a rich set of administration capabilities for both the Registry Administrator and the Governance Namespace Manager. The administrators may also copy and/or delete packages as requested by the Governance Namespace Managers or users via the Metadata Registry Administrators' GUI. Governance Namespace Managers and/or their Administrators review and modify the status on submission packages in their Governance Namespace. This task is accomplished through the Governance Namespace Administrators' GUI. The Registry Administrators (also known as Registry Operators or RegOps) are responsible for maintaining DSE content and administering users, systems, DSE usage courses, Governance Namespaces, and Import/Export processing. The DSE provides Registry Administrators with a GUI to:

- View and edit the Home Page
- Manage existing users and systems, and create new users and systems
- Manage the information on courses and training
- Track page counts

- Create and modify Governance Namespaces
- Assign roles
- Post resource submissions

Registry Administrators may also run various reports from the applicable metadata administration pages. These reports allow the Registry Administrators to get a list of information resources, the number of times a resource has been viewed, details on the number of times users have subscribed to resources, relationships between resources of the same or other Governance Namespaces, and details on the number of times resources are downloaded.

4.5 Interaction with Other CES and Structural Registries

As one capability within the Registry Services-sponsored suite of Core Enterprise Services (CES), the DSE is part of a tight federation that continually interacts to support NCDS objectives DoD-wide and beyond. Another essential requirement of the DSE is to support searches for structural metadata over the Internet and selected DoD networks. The Department's metadata is managed in numerous collections maintained by various organizations, and the collections are made searchable through the DSE's service interfaces. Users can then build simple or complex queries to search the metadata – concerning for example, geographic location, date of creation, or title. Users only need a common Web browser to search the DSE.

4.6 Operational Use Cases

The DSE provides support for users with a wide range of functions, skills, and sophistication. There are a number of common uses for the DSE, which will be discussed here. Each use case can have many variations depending on the lifecycle of the asset the metadata describes, the processes imposed by applicable governance bodies, and the role of the specific user with respect to the described asset and the DSE in general. There are also use cases associated with the management of the DSE itself.

Figure 11 illustrates the primary use cases of the DSE.

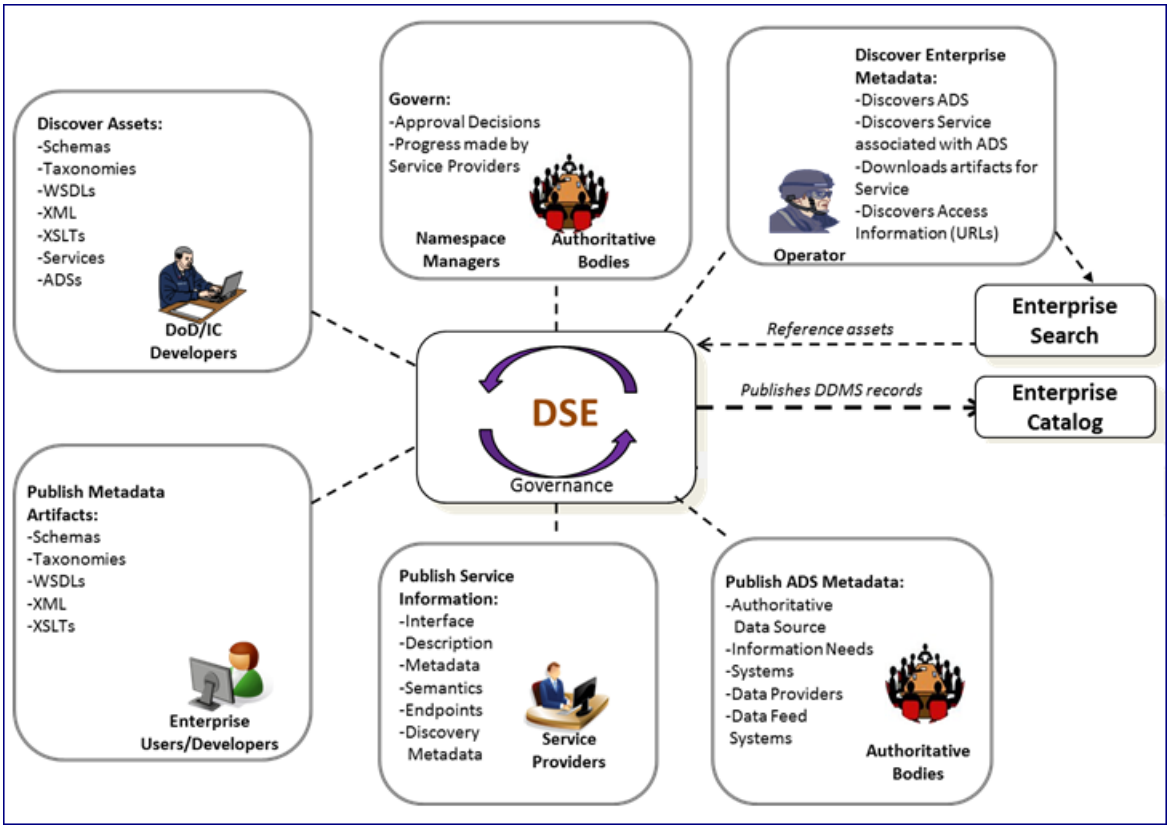

Figure 11. Top-Level DSE Use Cases

Figure 12 depicts a typical workflow that the DSE is intended to support with respect to publishing structural and semantic metadata. In these steps, the user, in this case a developer, builds a new metadata artifact (resource) that reflects necessary structure or semantics of importance and identifies a Governance Namespace in which it is believed the new resource should be registered. The developer coordinates with the Governance Namespace Manager to ensure appropriate placement of the resource. Upon registration, the Registry Administrator approves the submission causing the resource, which could describe concepts such as a service interface structure or standard terms for the namespace, to become visible. Following this approval, the metadata resource becomes visible within the specified Governance Namespace. Depending on the type of metadata artifact, components within the artifact can be parsed and made available as a set of searchable resources within the DSE. The Namespace Manager, now having CM authority over the resource(s), can modify the resource(s) status (Developmental, Operational, Deprecated, or Retired) as appropriate. The Governance Namespace Manager is responsible for ensuring that the status of the information resource(s) within the Governance Namespace reflects an accurate status based on communications from developers and other users.

As can be gleaned from the workflow, the most appropriate point of contact for a particular information resource is not necessarily the Governance Namespace Manager but the submitter of the information resource. Submitter information is always maintained with the information resource to encourage trust and to maintain a contact for ascertaining the pedigree of the information resource.

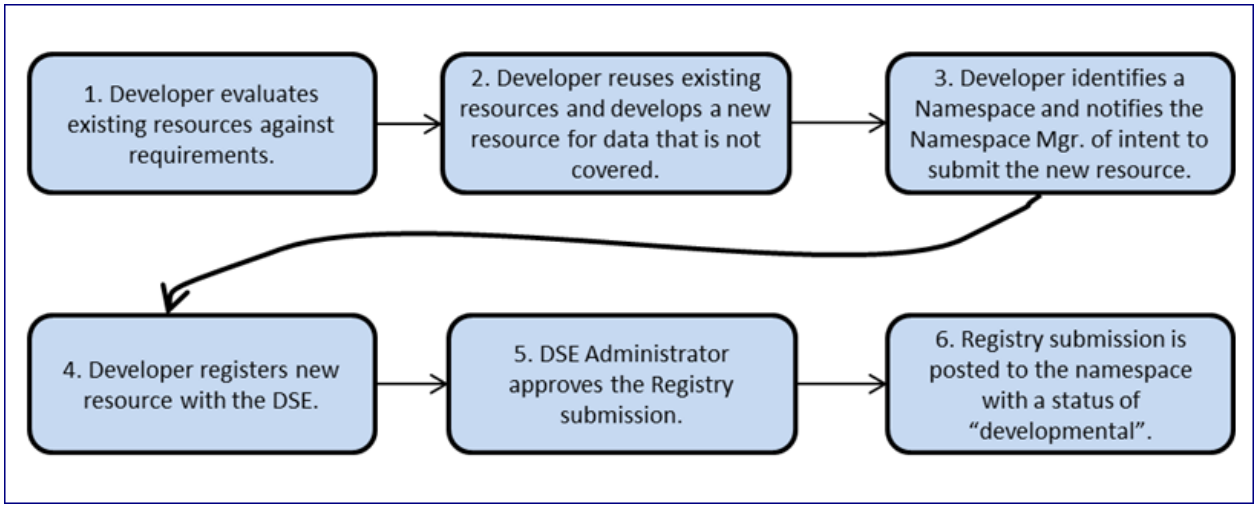

Figure 12. Publish Data Artifacts

Generally, the responsibilities of the various roles in this workflow are as follows:

- Users ensure that their submissions do not duplicate existing resources. They review existing resources for accuracy, currency, and applicability. They are responsible for coordinating with Governance Namespace Managers to ensure that the status of their resources is accurately reflected in the DSE.
- Registry Administrators review and approve all submissions. They also copy and/or delete resources as requested by users or Governance Namespace Managers.
- Governance Namespace Managers/Administrators review resources within their Governance Namespace, change the status as appropriate in coordination with users, and thereafter manage their resources in accordance with policies and processes endorsed by their governance authority.

Other aspects of metadata management revolve around the configuration control responsibilities associated with Governance Namespace Managers and Administrators. Governance Namespace Managers are responsible, to ensure configuration management of the assets within the namespace they are managing occurs.

4.6.1 Discover Assets

Figure 13 depicts a typical workflow that the DSE is intended to support with respect to publishing structural and semantic metadata. Many alternative flows for discovering assets are possible. In some cases the descriptive metadata stored in the DSE is sufficient for satisfying the user's needs; in other cases, the user will identify assets that can support the user, which the user will then download for analysis and use.

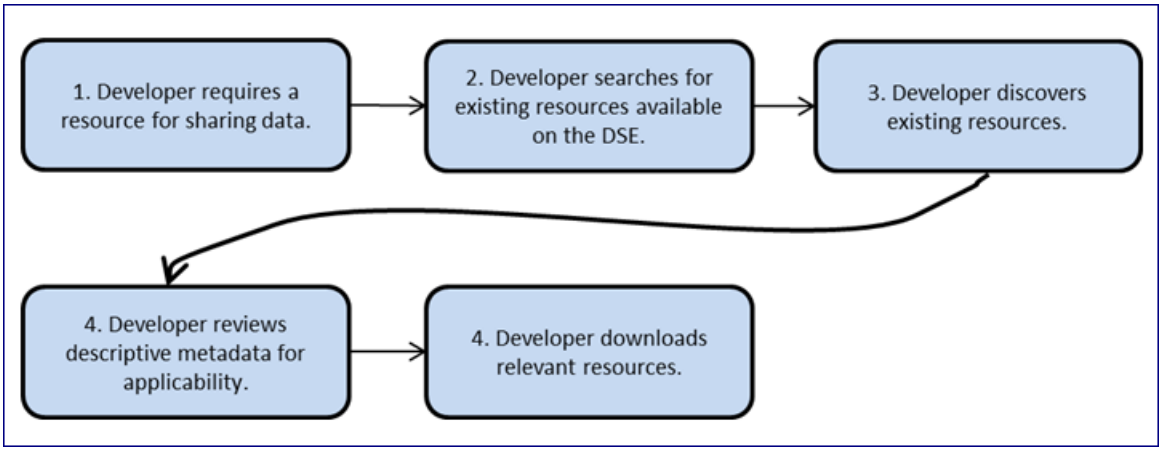

Figure 13. Discover Assets Workflow

4.6.1.1. Service and capability discovery

Use cases and patterns for the discovery of descriptive metadata about services or capabilities are explicitly supported by the DSE. Service consumers need to perform the following operations:

- **Static (Design-Time), User-Oriented Service Discovery**: Individual users and developers need to be able to browse, search, and inspect services and other entities via a Web-based GUI.

- **Dynamic (Run-Time) Service Discovery**: Service consumers may also need to discover services at run-time, using an inquiry web service interface provided by the registry. Dynamic inquiries are crucial to achieve location transparency of services, which allows service consumers to connect to a provider even if the service's location, or even transport, has changed.

- **Persistent Service Discovery**: In some cases a service consumer may want to be kept up to date on certain information entities published in a registry. A consumer should be able to advertise such a need by subscribing to changes of the interested discovery entities, and get notified of such changes near real time. This document will not discuss *persistent* service discovery in detail.

Regardless of the mechanisms, the service inquiries should have enough expressive power to support not just name and identifier based queries, but also complex query constraints based on arbitrary metadata. The following are some usage scenarios:

- **Joint Capability Area-Based Discovery**: Discover all services that are related to a given Joint Capability Area (e.g. Command and Control, Net-Centric, Logistics, etc.).

- **Capability-Based Discovery**: Discover all services that support a particular mission function, as defined by the SR&G endorsed taxonomy. (e.g. Decision Support, Search, etc.).

- **Data-Based Discovery**: Discover all services that support a certain type of data, such as moving indicator data.

- **Lifecycle-Based Discovery**: Discover all services in a given lifecycle (e.g. Developmental, Operational, etc.).

4.6.1.2. Static (Design-Time) Service Discovery

The static discovery of a service offer entity is done at design time by a human user, such as a software developer, by way of a Web-based GUI to discover the service offer and access points in the registry. For example, a software developer may need to discover a web service in order to bind to it from within a software module.

4.6.1.3. Dynamic (Run-Time) Service Discovery

In run-time service discovery, an application essentially performs the same operations that a human user performs for manual service discovery, except that the operations are invoked in an automatic manner. The automated discovery of service information entities is performed by an application that discovers, and subsequently invokes, the service described by the service information entities for purposes such as service composition (the dynamic invocation of one or more other services by a given service) and location transparency (enabling service consumers to connect to an access point even if the access location or transport has changed).

All of the usage examples described above for manual service discovery can also apply to automated service discovery. However, in dynamic service discovery, queries will more likely be focused around taxonomies that indicate technical and data-related capabilities rather than functional or human-readable metadata.

Dynamic service discovery is important when an implementation's access point might change. In this scenario, it is critical that the access point be automatically *rediscovered* (rather than using a cached version) on a regular basis, potentially each time it is accessed.

In a service-oriented environment, dynamic service discovery can help make system-to-system communications more loosely coupled by eliminating hardwired connections.

4.6.1.3.1. <u>Dynamic (Run-Time) Discovery Based Binding</u>

The discovery of services enables an important software feature called Dynamic Binding, sometimes also called Dynamic Invocation. Dynamic binding in modern object-oriented (OO) programming languages refers to the process where the client code invoking a method binds the actual implementation of the method dynamically at runtime. In the web services world, dynamic binding has been used to describe the process of binding the service client application to the service provider agent at runtime. The DSE plays a central role in this process by allowing the client application to do runtime discovery for dynamic binding. Dynamic binding is a showcase of the dynamic nature of SOA and the importance of Service Discovery as an essential component of a SOA infrastructure.

The common pattern of using dynamic binding is as follows:

1. The service consumer agent developer uses static discovery at design time to find the ideal service it needs to consume. The service consumer agent developer retrieves the technical information, that is, the service contract offered by the service offer. The consumer application is designed using this contract. This constitutes the compile time

binding, the binding of the consumer agent logic to one particular technical interface. The service consumer agent developer also develops runtime discovery code ready to use at runtime to find the access point of the desired service provider agent.

2. At runtime, the service consumer agent goes to the DSE again and looks for a service access point that conforms to the contract that the consumer agent is designed to use. The service consumer agent then actually binds to a service provider agent through the access point using the newly discovered access point information and possibly some other technical information, such as protocol, transport and security related information.

Runtime Discovery of a Known Service Access Point by Identifier

At design time, the consumer agent is bound to the provider agent. At runtime the consumer agent only needs to bind to the access point. This is the simplest case in which the developer on the consumer side already chose the provider agent at design time. The service provider agent's unique identifier/key is already known and "bound" to the consumer application. At runtime, the consumer application simply goes to the DSE to retrieve the up-to-date access point of the target service provider agent, and possibly caches the endpoint information for a certain period of time. The discovery capability used is the **get** capability. This usage is very similar to a DNS client lookup of the IP address of a host name. This approach has been used in Business to Business (B2B) integration scenarios where business software needs to do runtime lookup of their business partners' service endpoints.

Runtime Discovery of a Technically Compatible Service Access Point

At design time, the service consumer agent binds to the technical *interface* of a service offer. The consumer agent needs to dynamically bind to the service provider agent at runtime. This is a more complicated case than simply locating a specific provider agent's access point. The service consumer agent is designed and programmed to consume a service offer with some particular technical interface, such as represented by a WSDL portType. At runtime, the service consumer agent goes to the DSE to discover any technically compatible service provider agents and their access points by using variety of matching attributes. Upon receiving the search results from the DSE, the service consumer agent chooses to bind to the desired service agent's access point on-the-fly. The consumer agent needs to decide the criteria for a *desired* service provider agent and access point.

The interesting aspects of this use case come from the flexibility the consumer agent has over the following:

1. The kind of service discovery query it can send to the DSE at runtime
2. How the consumer application uses the results of the search

Some consumer agents have been designed and programmed to send a query simply to find all service producer agent access points that conform to one particular service specification (for example, WSDL portType). More often, however, consumer agents may need to send a query that contains a filter on other metadata attributes. For example, one consumer agent may be interested in locating access points that conform to the Enterprise Search Service Specification, offered by the Department of Air Force, and serving data about old Soviet Sukhoi fighter jets.

An example of such a consumer agent would be an aggregated Web Portal for content search, where the user interface allows the user to specify a set of metadata attributes.

Consumer agents may also restrict their search using service access point level technical categorizations, such as QoS information, hosting location information, and security policy requirement of the service access point (for example, whether a service access point requires authentication and what type of authentication is required). One particular usage of such runtime discovery is for load-balancing and fault-tolerant fail-over to a compatible service. For example, the Registry Services Service Security policy enforcement message handler can use the runtime discovered *authoritative* Policy Decision Service access points for fail-over purpose: if one is not available, try next one, thus achieving better service continuity and availability.

Runtime Discovery of a Service Access Point That May Be Technically Incompatible

Some service consumers may not want to bind to a single technical interface. Instead, the consumer agent's business logic is more concerned about the functionalities and semantics of a target service's operations to be used. The service consumer agent then relies on some transformation service to serve as an adapter to bind to the runtime discovered service access point. The necessary SOA infrastructure to this kind of *just-in-time* integration is supported by the DSE but involves design patterns and components beyond the scope of the DSE and UDDI. It is beyond the scope of this document to address these issues in full detail.

4.6.1.4. Subscription Based Service Discovery

Also known as persistent service discovery, service subscription involves a subscription by a service consumer or service consumer agent to certain information entities in the registry and subsequent near-real-time notifications upon changes to that service entity.

4.6.2 Govern

Figure 14 depicts a consolidated workflow that the DSE can support with respect to publishing structural and semantic metadata. The actions available to a given user depend on the roles they have across the variety of metadata assets and the governance actions that are needed at that time.

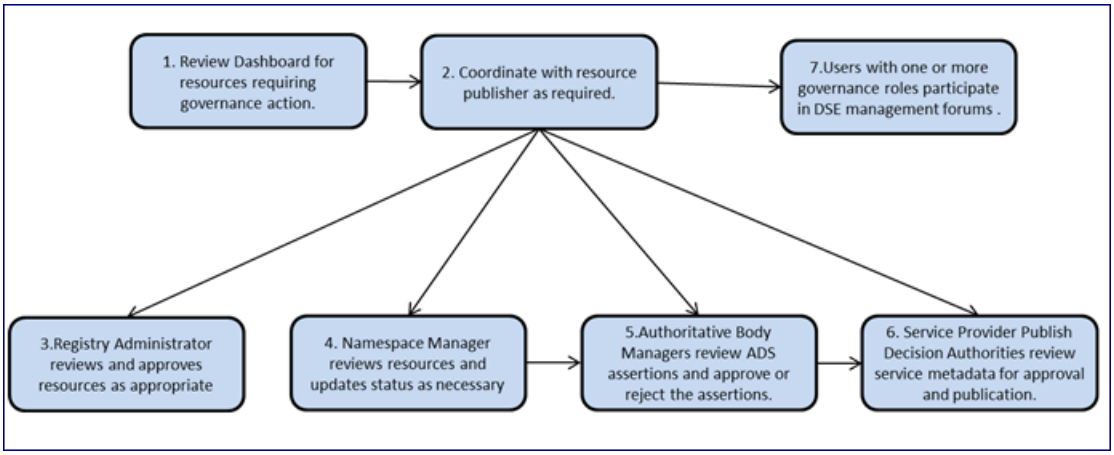

Figure 14. Govern DSE and Metadata Resources

4.6.3 Discover Enterprise Metadata

Figure 15 depicts user activity that the DSE can support with respect to discovering metadata and other assets in repositories and registries beyond the DSE. The user can identify information needed from the Enterprise Search capability that may not be present in the DSE. Upon receiving results, the user is responsible for accessing the identified resources, and adapting them for reuse as appropriate.

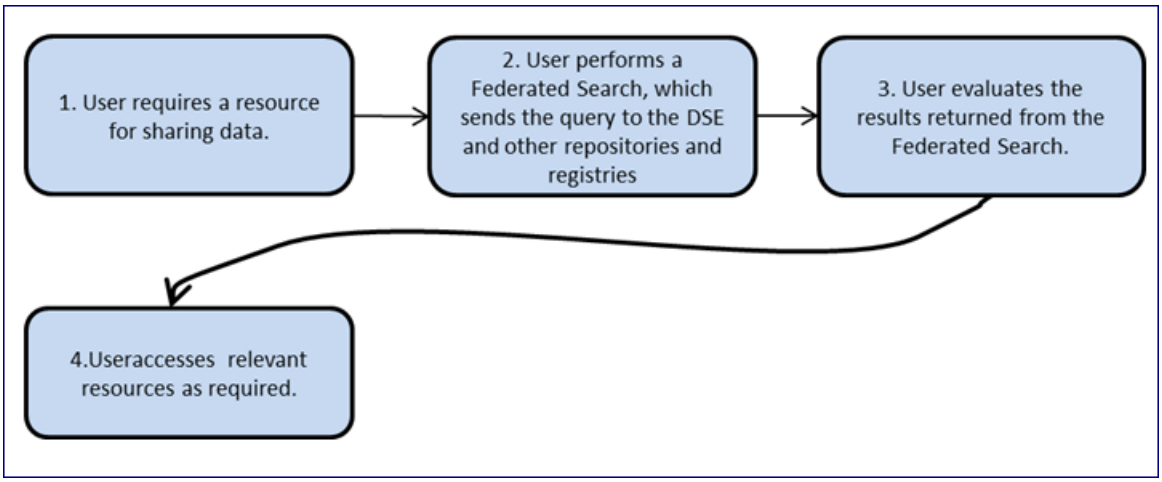

Figure 15. Federated or Enterprise Search

4.6.4 Publish Authoritative Data Source Metadata

In order to support the Discovery capabilities above there is a need to support a robust registration and governance model that will allow users to populate the DSE with information sufficient for authorizing bodies to make informed decisions about proposed authoritative data sources. Figure 16 illustrates the core functionality of the DSE from an Authoritative Data Source registration and approval perspective.

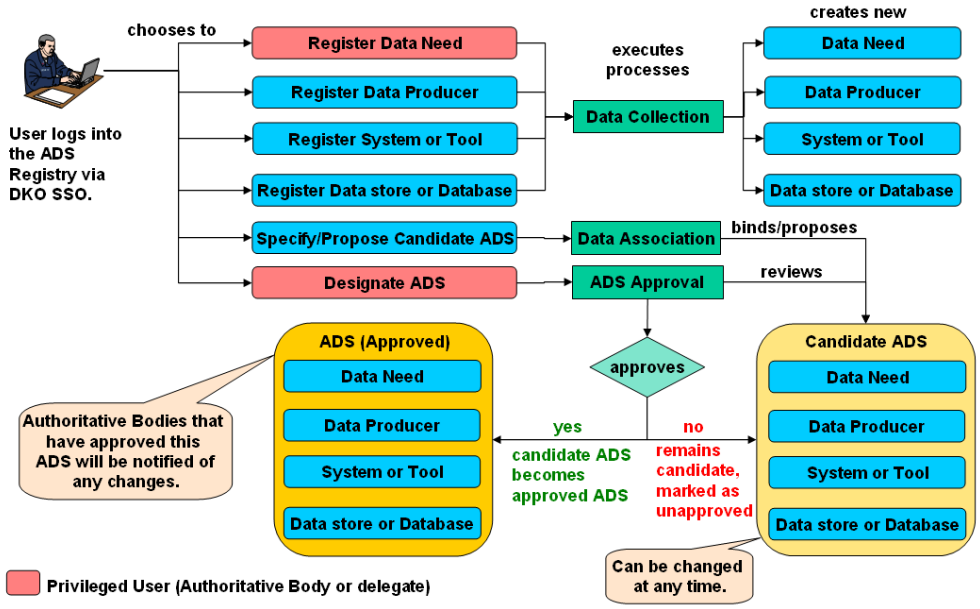

Figure 16. Authoritative Data Source Registration, Proposal, and Approval Process

As illustrated in Figure 16, there are three main ADS processes supported by the DSE:

1) The Data Collection Process – This process supports the registration of new data needs, data producers, systems or tools, and data stores or databases.

2) The Data Association Process – This process supports the binding of data needs, data producers, systems and tools, and data stores and databases and other data feeds together to form a candidate authoritative data source. (Recall that a data need is associated with an authoritative body which determines which authoritative body a candidate authoritative data source is required to be approved by in order to become an approved authoritative data source.) The data association process also allows for the binding of data needs to authoritative bodies.

3) The ADS Approval Process – This process provides a simple way for an authoritative body or a designee thereof to view the data needs and proposed authoritative data sources and make approval decisions based on the information provided. Once an approval occurs a proposed authoritative data source becomes and approved authoritative data source.

Between these three processes the DSE will support the definition of data needs and the association of those needs to an authoritative body, as well as the registration of any data source and allow it to be proposed as authoritative for a given need. The approval process that candidate authoritative data sources are subject to provides thorough traceability so that users seeking authoritative data can be assured that designated authoritative data sources are reliable, maintained, and can be trusted and relied upon for use in their specified operational contexts.

4.6.5 Publish Service Information

For any of the supported information resource types, a user should be able to discover, retrieve, understand, rate, and relate existing information resources as well as register new information resources. Analysts should be able to discover, retrieve, understand, rate, and relate information resources and Governance Namespace Managers should be able to rate, relate, and register information resources. The DSE provides capabilities that enable each of these use cases. Figure 17 illustrates the interactive (human in the loop) service publication use case. Additionally, the DSE supports the service publication WSDL pattern.

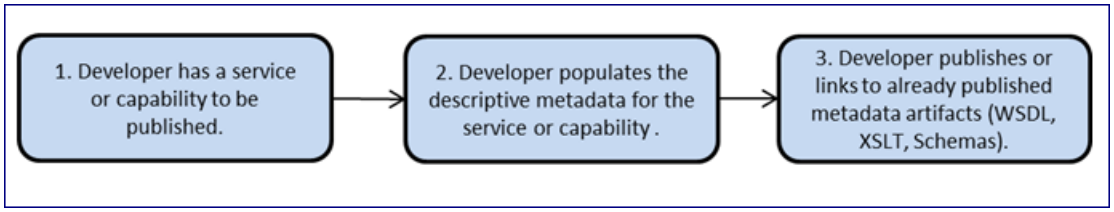

Figure 17. Interactive Service Publication

4.6.5.1. Taxonomy Management

The UDDI v3 specification does not address a standardized taxonomy format and management system. While different vendors have implemented varying value-added features to support taxonomy management within a UDDI registry, this document will not focus on any specific vendor implementation of taxonomy support. Instead, this document will lay out general guidelines for taxonomy management and will defer to the OASIS UDDI Technical Note, "Providing a Value Set for Use in UDDI Version 3", for guidance on publishing taxonomies in a UDDI registry [UDDI-TAX]. The Service Registry will implement taxonomies as defined and endorsed by the SR&G WG.

The Service Discovery component provides two RECOMMENDED taxonomy management best practices:

- **Importing externally managed taxonomy data**: Because a taxonomy may contain many values, manually inputting them into a registry is often not desirable. Furthermore, the taxonomy may need to be reused and shared among multiple registries and should therefore be put in an interoperable exchange format. It may be more convenient and efficient to store the master copy of the taxonomy data in an external store (for example, an XML file or a metadata repository) and then import it into the registry.

- **Categorizing the taxonomies**: As discussed in the Categorization section of this document it is a best practice to associate the taxonomy itself with a number of categories such as requirement, dimension, and applicability.

The process of publishing taxonomies in UDDI is based on the OASIS UDDI Technical Note, "Providing a Value Set For Use in UDDI Version 3" [UDDI-TAX]. UDDI refers to taxonomies as being checked or unchecked. Checked taxonomies require that all references to the taxonomy be validated against a set of valid values, while unchecked taxonomies do not have this requirement. A checked taxonomy must have an associated validation service that performs value checking each time an attempt is made to save data that contains a reference to the

taxonomy (for example, to ensure that the taxonomy value that is referred actually exists within the published taxonomy).

To publish unchecked taxonomies, all that is required is the publishing of a tModel that represents the taxonomy. Publishing checked taxonomies is more complicated and requires the publishing of a validation service for the taxonomy. Categorization of the taxonomies themselves in UDDI can be done simply by providing a categoryBag element for the tModel that represents a taxonomy.

4.6.6 Runtime (Machine-to-Machine Interactions)

Concerning machine-to-machine vice human usage, the following interactions are anticipated:

- With other Core Enterprise Service metadata engines such as the Service Registry, Federated Search Content Catalogs, etc.
- With structural metadata repositories and tools fielded outside the Registry Services Program
- With Test and Evaluation environments put up by various communities.

The diagram below illustrates the primary set of use cases that the DSE is intended to realize for the purposes of automated runtime access.

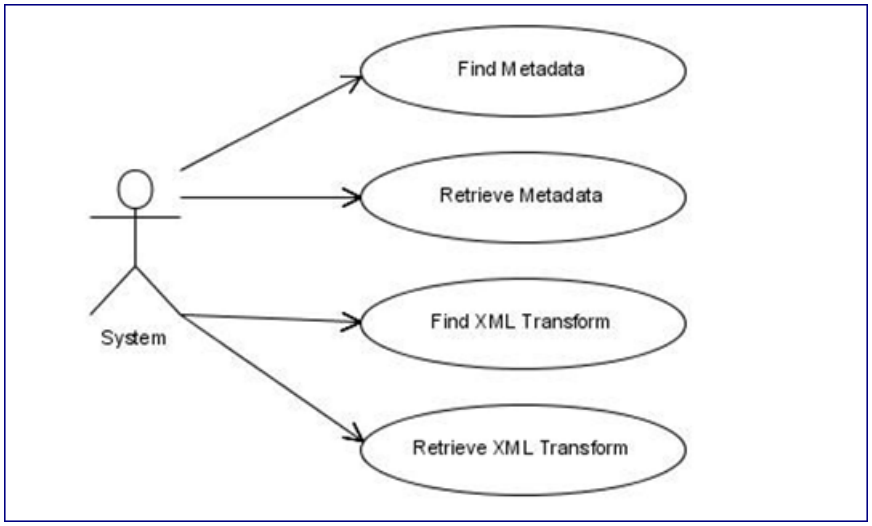

Figure 18. Runtime Use Cases

In this diagram, the user is a system that needs to find and retrieve information resources such as XML schema documents, XSL transformation documents, taxonomies, or domain value documents for use at execution time. This might be for sharing information between systems or populating user interfaces. The DSE supports these use cases by providing web services for machine-to-machine communication. Detailed information on using the DSE's web services, see the DSE's web services documentation.

5. Summary

The DSE simplifies the publication and discovery of data services and facilitates information sharing across the DoD, providing an enhanced web dashboard, unifying access to many service capabilities and simplifying access for users and providing SSO to enterprise capabilities. Through the continued progress of Net-Centric Publisher, Service Discovery, the DoD Metadata Registry, and the Enterprise Authoritative Data Source Registry, DSE's success has shown successful rapid integration of existing technologies, combining them in such a way to provide common interfaces and abstracting complex logic and detail.

As we move forward into the future, the DSE can easily evolve to provide more integration capabilities – moving from publication, governance, and discovery, to facilitating integration with other enterprise components. This may provide an incentive to service providers to register their services if security implementation can be provided as part of the process.

As we move forward into the DoD Cloud, the components of the DSE could become critical aspects of PaaS. In various cloud environments (private, public, hybrid models), the registry becomes the key element for facilitating registration, discovery, and integration on all platforms. What becomes crucial in this environment is the ability for the registry to be distributed across a cluster, consistent, and highly reliable, providing Data as a Service (DaaS) and facilitating integration between clients of various platforms and devices, and applications, widgets, and services.

The DSE can be enhanced to offer its hosted data collections as a DaaS component of the DoD Cloud infrastructure. Common re-usable components, such as WSDLs, schemas, and taxonomies, could be made available to other the data producers across the enterprise for incorporation into their solutions. Providing such functionality could enable cost efficiencies across the enterprise application developer communities. Individual capability developers will no longer be required to define, configure, and maintain a data environment for that capability's use. Leveraging shared common data will decrease development cost, deployment time, and security assessment effort, and will also reduce shared support costs for individual data solutions, and DSE would increase effectiveness of the capability by introducing value-added layers to the data, such as data processing, output or mediation across vocabularies.

In the future, DSE can leverage the service registry as a Widget and App Store repository, in order to enable the advertisement and delivery of enterprise capabilities in support of activities surrounding widget frameworks and applications storefronts. The DSE already enables users to capture rich descriptive information about services that can be useful both to the human user searching for a particular genre of services or to a machine searching for a service with a very specific API. DSE offers bi-directional federation capability, which will allow it to be the single authoritative source of services, while promoting distributed access to hosted services. These existing capabilities solidly position the DSE capability to easily become the backbone for DoD's Enterprise App Store for both the enterprise-user desktop, and for mobile devices. This will realize cost savings by avoiding separate acquisition costs for multiple proprietary app store infrastructure that the Department pursues (e.g., Apple, Droid, etc.). Furthermore, new component types can be supported with new taxonomies within the DSE, which can be done without any software development process or expense. It will also promote greater effectiveness

of the Data and Service strategy goals, by maintaining a single point of access for all service types. A single, consolidated authoritative repository will increase visibility into potentially duplicate capabilities that can be consolidated, thereby realizing cost savings.

As the DSE looks towards the future, we should continue to move forward in providing more holistic and integrated functionality, as it will enhance data discoverability, accessibility, and currency while making data and service integration easier.

Appendix A: Acronym List

Acronym	Definition
API	Application Programming Interface
ATO	Advanced Technology Office
B2B	Business to Business
C2	Command & Control
C4I	Command & Control, Communications, Computers and Intelligence
CoI	Community of Interest
COMINT	Communications Intelligence
CONOPS	Concept of Operations
COTS	Commercial Off The Shelf
CRM	Customer Relationship Management
DISA	Defense Information Systems Agency
DOA	Distributed Object Architecture
DoD	Department of Defense
DTD	Document Type Definition
ELINT	Electronic Intelligence
GOTS	Government Off The Shelf
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol, Secure
IDL	Interface Definition Language
J2EE	Java 2 Enterprise Edition
JMS	Java Message Service
LDAP	Lightweight Directory Access Protocol
OASIS	Organization for the Advancement of Structured Information Standards
OO	Object Oriented
PDA	Personal Digital Assistant
PKI	Public Key Infrastructure
QoP	Quality of Protection
QoS	Quality of Service
RO	Responsible Organization
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
UAV	Unmanned Aerial Vehicle
UDDI	Universal Description, Discovery, and Integration

URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WS	Web Services
WS-I	Web Services Interoperability (Organization)
WSDL	Web Services Definition Language
XML	Extensible Markup Language

Appendix B: References

[RFC 2119]	S. Bradner. Key words for use in RFCs to Indicate Requirement Levels. IETF RFC 2119, March 1997. http://www.ietf.org/rfc/rfc2119.txt
[UDDI-3]	UDDI Version 3.0 Specification, http://uddi.org/pubs/uddi-v3.00-published-20020719.htm
[UDDI-TAX]	"Providing a Value Set For Use in UDDI Version 3", http://oasis-open.org/committees/uddi-spec/doc/tn/uddi-spec-tc-tn-valuesetprovider.htm
[UDDI-WSDL]	"Using WSDL in a UDDI Registry, Version 2.0.2", UDDI Technical Note, June 31, 2004, http://www.oasis-open.org/committees/uddi-spec/doc/tn/uddi-spec-tc-tn-wsdl-v2.htm
[WSDL-1.1]	Web Services Description Language 1.1, March, 2001 http://www.w3.org/TR/wsdl
[WS-I-BP]	WS-Interoperability Basic Profile 1.0, Web Services Interoperability Organization, http://www.ws-i.org/Profiles/BasicProfile-1.0.html