



Army Information Architecture (AIA)

Version 4.1

Project Number: 0035

Prepared by:

Office of the Army Chief Information Officer
Architecture, Operations, Networks and Space Mission
Information Architecture Division
Arlington, VA 22202

and

CECOM Life Cycle Management Command
Software Engineering Center (SEC)
Army Net-Centric Data Strategy Center of Excellence (ANCDS CoE)
Aberdeen, MD 21005

Revision History

Revision	Date	Description of Change
1.0	23 Dec 2010	Draft Release
1.2	18 Feb 2011	Spiral 2 Pre-Release
2.0	28 Feb 2011	Spiral 2 Release
2.1	25 Mar 2011	Addition of implementation priorities and additional processes
3.0 Draft	27 May 2011	First Draft Release of Spiral 3
3.0 Draft 2	15 Jun 2011	Second Draft Release of Spiral 3; addresses comments received from wide review.
3.0	30 Jun 2011	Final Spiral 3 Release
3.1	30 Mar 2012	Addition of Cloud Computing Guidance. Expansion of Secured Availability Guidance. Incorporation of responses to deferred AIA 3.0 comments. General editorial changes and clean-up.
4.0	29 June 2012	Improved integration with ADF chapters; result was two new level 1 sections: "Data Asset Development and Management" and "Data Delivery and Use." Expanded Data Quality Guidance. Expanded guidance to include disadvantaged (DIL) computing environments. Expanded use/application of DoDAF. Clarified explanation of relationship to DoD IEA 2.0 to account for changes between DoD IEA 1.2 and 2.0. General editorial changes and clean-up.
4.1	07 Dec 2012	Added and revised business rules for coordination with <i>Rules for Cross-Cutting Capability (CCC) Information Exchange Specifications (IES) in Interface Specifications</i> [23]; Updated references to MDR and EADS to DSE 2.0.
4.1	30 Apr 2013	Updated to respond to Army Staffing Review comments. Changes did not impact version number.

Table of Contents

1.	Introduction	1
1.1	Net-centricity, Information Sharing, and the Future Army	1
1.2	Purpose	1
1.3	Scope	2
1.4	Objectives.....	4
1.5	Applicability.....	4
1.6	Assumptions.....	5
1.7	Audience/Stakeholders.....	5
1.8	Army Information Enterprise Priorities.....	6
1.9	Document Overview	7
1.9.1	Overview of AIA Document Organization	7
1.9.2	Overview of Principle and Business Rule Organization	9
1.9.3	AIA Relationship to Other Documents	10
2.	How Is This Document To Be Used?	14
2.1	Guidance	14
2.2	Compliance.....	14
3.	End-State Information Sharing Framework	16
4.	Disconnected, Intermittent, and Limited (DIL) Networks and Limited Capability CEs.....	19
5.	Global Principles and Business Rules	21
5.1	Data and Information	21
5.2	Data Exchange and Information Sharing	22
5.3	Information Sharing Governance and Guidance	23
6.	Data Asset Development and Management (DADM)	25
6.1	Foundation for Data Exchange and Information Sharing.....	25
6.2	Physical Structuring of Data.....	26
6.3	Data Model Guidance	28
6.4	Data Quality.....	30
6.5	Data Integration	31
7.	Data and Services Deployment (DSD)	33
7.1	Enabling Information Sharing and Usage.....	33
7.2	Data Exchange Planning and Implementation	35
7.2.1	Anticipated and Unanticipated Information Sharing	35
7.2.2	Information Exchange Specifications.....	35
7.3	Data Asset Deployment Planning and Implementation	37
7.3.1	Authoritative Data Sources (ADS)	38
7.3.2	Unstructured Data Assets	39
7.3.3	Cloud-based Data Assets.....	40
7.3.4	ERP-based Data Assets.....	42
7.4	Data Service Planning and Implementation	43
7.4.1	Data Services.....	43
7.4.2	Data Service Guidance	46
7.5	Interoperability Planning and Implementation	48
7.5.1	Master Data Management (MDM)	48
7.5.2	Community-based Information Sharing	49
7.5.3	Translation and Mediation	52
7.6	Discovery and Accessibility Planning and Implementation	53
7.6.1	Metadata Management	53
7.6.2	Registration.....	55

8.	Data Delivery and Use (DDU)	57
8.1	Enabling Data Delivery and Use	57
8.2	Information Requirements Traceability	57
8.3	Dashboards and Portals	58
8.4	Business Intelligence	59
9.	Secured Availability (SA)	60
9.1	Enabling Secured Availability and Access	60
9.2	Information Assurance (IA)	61
9.3	Data Security	62
9.3.1	Classification	62
9.3.2	Media and Devices	63
9.3.3	Encryption	63
9.3.4	Transfer	63
9.3.5	Disposal	64
9.4	Data Service Security	64
10.	Governance	67
10.1	Chief Data Officer (CDO)	67
10.2	Army Data Board (ADB)	68

Figures

Figure 1:	Scope of an Information Architecture	3
Figure 2:	Document Organization Overview	8
Figure 3:	Organization of Principles and Business Rules	9
Figure 4:	AIA Document Context	10
Figure 5:	Relationship of AIA to Other Document Products	12
Figure 6:	End-State Information Sharing Framework (Simple Form)	16
Figure 7:	Army Information Architecture (AIA) End-State Information Sharing Framework	17
Figure 8:	Examples of Data Assets	25
Figure 9:	Service Reuse and Development Flow Chart	45
Figure 10:	Metadata Categories	54
Figure 11:	Dashboard and Portals	58
Figure 12:	COE Implementation Plan Technical Reference Model (TRM)	99
Figure 13:	Data Standard Classifications	100
Figure 14:	Process of Establishing an Authoritative Data Source in DSE 2.0	111
Figure 15:	Lifecycle of Technical Specifications	113
Figure 16:	Rainmaker DDF (Conceptual)	115
Figure 17:	Data Service Development Process	122

Tables

Table 1:	Relationship of AIA to DoD Net-Centric Data Sharing Objectives	13
Table 2:	Relationship of AIA to DoD Reference Architecture Elements	13
Table 3:	Standards Adoption Priority	24
Table 4:	Levels of Interoperation Description	51
Table 5:	AIA Companion Products	97
Table 6:	Foundation Data Standards	101
Table 7:	Infrastructure Data Standards	101
Table 8:	Data Exchange Standards	102
Table 9:	Data Access Standards	104

Table 10: Consumer-Accessible Data Services.....	106
Table 11: Data Management Services	106
Table 12: CDR Data Service Descriptions.....	107
Table 13: DSL-A Data Service Descriptions.....	108
Table 14: Data Service Development Process Description.....	122

Appendices

Appendix A References	69
Appendix B Acronyms and Definitions	75
B.1 Acronyms and Abbreviations	75
B.2 Definitions.....	80
Appendix C List of Principles and Business Rules	83
Appendix D Relationship to Other Data Strategy Products.....	97
Appendix E Catalog of Data Standards.....	99
E.1 Data Standards.....	99
E.2 Foundation Data Standards.....	101
E.3 Infrastructure Data Standards.....	101
E.4 Domain-Specific-Information Data Exchange Standards.....	102
E.5 Data Access Standards	104
Appendix F Data Services.....	106
F.1 Data Service Family Descriptions	106
F.2 Data Service Descriptions.....	107
Appendix G Processes and Activity Models	110
G.1 Community of Interest Processes	110
G.2 Data Planning Processes.....	111
G.2.1 Authoritative Data Source Processes.....	111
G.2.2 Information Exchange Specification Processes.....	112
G.2.3 Unstructured Data Processes.....	113
G.2.4 Data Model Planning and Organization	116
G.2.5 Interoperability Mapping, Translation, and Mediation Processes	119
G.3 Service Planning Processes	121
G.3.1 Data Service Development Process.....	121
G.4 Migrating Data to a Cloud Computing Environment	123
G.4.1 Introduction	123
G.4.2 Cloud Data Deployment	124

Executive Summary

The Department of Defense (DoD) Information Enterprise Architecture (IEA) [1] [2] [3] provides a foundation to support DoD transformation to net-centric operations. The U.S. Army's Common Operating Environment (COE) Architecture [4] provides guidance to program managers and solution developers in the selection and use of approved computing technologies and standards to foster integration and interoperability. This document, the Army Information Architecture (AIA), provides the design and development guidance needed by Army personnel, from Communities of Interest (COIs) to Program Managers (PMs) to developers, to create information systems that meet DoD and Army net-centricity and information sharing objectives. The AIA:

- Complements the DoD IEA by providing “bottom-up guidance in the form of Army-specific principles, business rules, and processes that govern data and data service design, development, and deployment.
- Augments the COE by identifying the data and service standards that complement the established COE computing technologies and establishing the processes and governance for Army data and data service deployment that facilitate information sharing.

The AIA provides the foundation to accelerate Army transformation to net-centric information sharing in two (2) ways. The first is as design and development guidance for enabling information sharing. The second is as a set of compliance requirements for assessing the level to which systems meet net-centric information sharing objectives.

The AIA presents an end-state information sharing framework that presents the key concepts involved in net-centric information sharing and their interrelationships.

The primary content of the AIA is a collection of principles and business rules that are organized around and address the following topics: data asset development and management; data and services deployment; data delivery and use; and secured availability. A principle is a generalized statement of position that is accepted as true or valid, and often reflects values, beliefs, or convictions on the “right” or “best” way to do or achieve a result or fulfill a mission. Business rules are by-products of principles; they are recommendations, requirements, guidelines, directives, stipulations, or imperatives that assert what shall/should be done to meet or implement the principle. For example, the principle: “Effective information sharing is based on clear, unambiguous, and consistent management of structured data” is supported by the business rule “A schema shall be developed and maintained for each structured data asset (e.g., database, data service interface, or message format).”

Additional information that supports the primary content of the AIA, such as a description of a data service development process, is provided in the appendices.

The intended users of this document are Headquarters, Department of the Army (HQDA), Data Stewards, Functional Data Managers (FDMs), PMs, Program Executive Offices (PEOs), architects, and developers. The document is used in the planning, development, and implementation of solutions and/or systems that support effective information sharing among the Warfighter and other agents throughout the Army.

The Army Information Architecture is approved for immediate use. My Point of contact for this document is Mr. Cliff Daus, Division Chief for Information Architecture Division at (571) 256-8953 or cliff.a.daus.civ@mail.mil.

E-Signed by BLOHM, GARY W. 1228949589
VERIFIED BY: [Signature]
BLOHM, GARY W. 1228949589

GARY W. BLOHM
Army Chief Data Officer
Director, Army Architecture
Integration Center

1. Introduction

1.1 Net-centricity, Information Sharing, and the Future Army

The Department of Defense (DoD) will conduct future business operations, warfare, and enterprise management using Information Technology (IT) that provides an assured, dynamic, secure, and shared information environment – an environment that provides access to trusted information for all users, based on need, independent of time and place. Net-centric warfare translates information superiority into combat power by enabling knowledgeable entities in the enterprise and battlespace to effectively share information.

To achieve information superiority and a common operational picture, information must flow to those who need it and empower not only the commanders and other decision makers, but also individual warfighters and support personnel. Sharing a common situational understanding of the battlespace and enterprise is essential for military and support operations, but building the IT infrastructure to enable and provide that common understanding is not a simple process. A top-down strategic vision for net-centric interoperability is needed that is coupled with bottom-up tactical guidance that moves systems, step-by-step, toward that vision.

The heart of net-centric interoperability is information sharing: the ability to rapidly and securely send, find, access, and use information is the key to an agile Army and DoD Enterprise. The vision for information sharing in the Army, as stated by the Secretary of the Army, is:

“Data is a strategic asset and must be managed as such. [The Army’s] goal is to create and support a network-enabled environment that gives decision makers access to data in a timely and secure manner.” [5]

To achieve this vision of information sharing, the Army needs to transition from an environment of information stovepipes to one of globally-accessible, secure, re-usable information.

1.2 Purpose

The purpose of this document is to provide the specific information sharing and data exchange guidance and compliance requirements that will enable Army stakeholders to envision, design, develop, deploy, and use information systems that:

- (1) are consistent, comprehensive, compatible, and integrated in their ability to share information across the Army, and
- (2) realize the DoD information sharing vision and meet the Army information sharing objectives.

The AIA is the starting point for understanding the Army information sharing and data exchange guidance and compliance requirements. The end-state information sharing framework is presented in Section 3; the framework reflects the perspective of, and serves the purposes, of the AIA.

This document augments, extends, and complements the DoD Information Enterprise Architecture (DoD IEA) [1] [2] [3] and the Army’s Common Operating Environment (COE) Architecture [4] to meet Army-specific requirements and provide information sharing and data exchange guidance. This document is a *Reference Architecture* (RA) as defined the DoD Reference Architecture Description [6]; a description of how the Army Information Architecture (AIA) meets the requirements of the Reference Architecture Description is presented in Section 1.9.3.

The DoD IEA is the starting point for establishing requirements for compliance with the DoD's information sharing strategies. The DoD IEA:

"...addresses the concepts, strategies, goals and objectives related to the IE <Information Enterprise> and provides a common, enterprise foundation to guide and inform IT planning, investment, acquisition and operational decisions in achieving the IE vision. It describes the IE capabilities that enable DoD operations by establishing the activities, rules and services involved in providing the IE capabilities." ([2], p. 1.)

The COE is the starting point for choosing approved technologies for the development of interoperable Army information system solutions. The COE:

"...is an approved set of computing technologies and standards that enable secure and interoperable applications to be developed and executed rapidly across a variety of computing environments (i.e., server(s), client, mobile, sensors and platform)." ([4], p. 7)

This document adopts the COE as the definition of the platform(s) that is the foundation upon which information sharing objectives will be met.

1.3 Scope

The scope of the AIA is information sharing guidance for data that is stored within, and that moves throughout, Army networks, including all COE Computing Environments (CE; see COE [4]). This includes data that moves into and out of Army networks through interfaces to external systems (e.g., DoD, Joint systems), though the scope of this document "stops" at those interfaces.

The scope of information sharing guidance presented in this document is:

- the definition of priorities, principles, and business rules, and
- the identification of the processes, standards, patterns, and implementation aids

that guide, govern, or support the design and implementation of data format, data service, and data management features of Army IT systems and facilitate information sharing among those systems.

The material in this document presents or cites only the requirements, standards, processes, and actions that must be met to achieve Joint, DoD and Army information sharing objectives. It does not constitute an exhaustive account of *all* requirements, standards, or actions that must be met by system development efforts. For example, while the guidance in this document overlaps with conventional data management practices, many data management practices (e.g., archiving, replication, wiping) are out of scope of this document.

The scope of this document focuses on high-level, general, information sharing guidance that applicable enterprise-wide and overarches specific topic areas like data quality or cloud computing. The AIA introduces some specific topic areas, but does not provide complete guidance necessary for those topic areas. Rather, the AIA is intended to be the entry point to the topics and reference other guidance documentation that drills-down into the details necessary for fully addressing those topic areas.

The scope of the AIA as an *information architecture* is illustrated in Figure 1. The cube in the diagram shows three (3) architectural views of the underlying system. The Technology Architecture view shows how physical components like computers and routers are connected. The Software Architecture view shows how software components are structured and how they

interact; the software is dependent on hardware in the technology view, but can, to a large degree, be considered and architected independently of the technology. The Business Architecture view shows the business processes of the system; the Technology Architecture and Software Architecture must ultimately support the business processes in the Business Architecture.

The Information Architecture cuts across all three of these views, focusing specifically on data, data access via services, and how data moves to support the information sharing requirements of the business processes. Examples of each type of architecture are depicted in Figure 1.

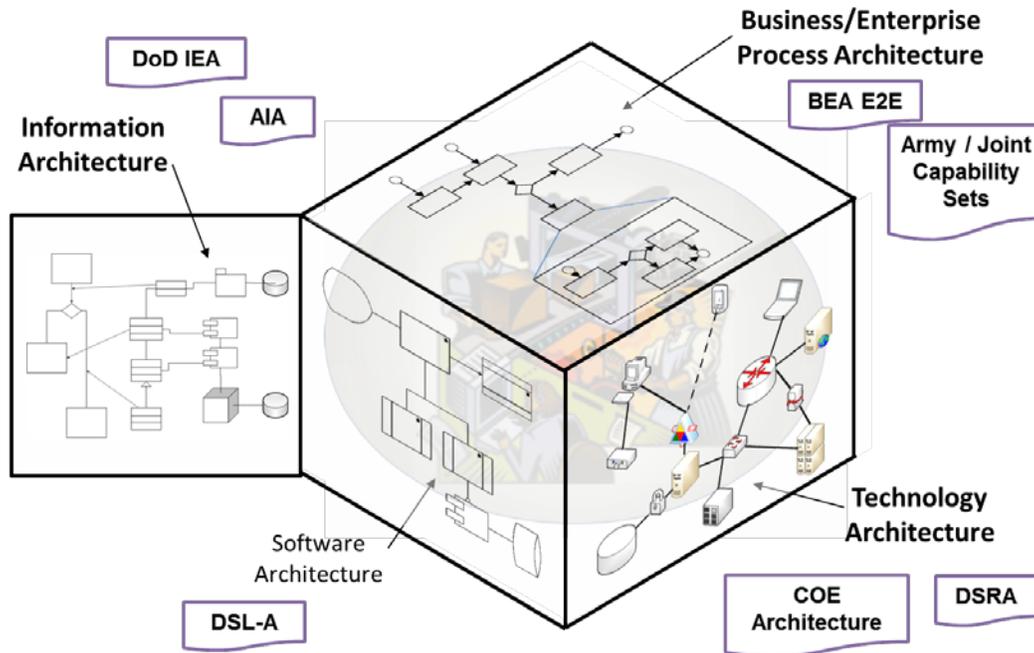


Figure 1: Scope of an Information Architecture

The scope of this document includes the criteria and other information needed to assess compliance to the business rules specified in this document and the DoD IEA.

The scope of this document does not include:

- physical technology that processes data, e.g., equipment and networks, which is covered by the COE [4];
- low-level utility software technology, e.g., operating systems, also covered by COE [4];
- the business processes that use data or data requirements for mission fulfillment;
- prioritization of data, data assets, or data services for adoption or application of the AIA;
- detailed planning for the adoption or application of AIA;
- software development lifecycle processes;
- data flow or data management within the boundaries of an application or single, bounded system;
- project execution, including costs and schedules; and
- very large scale system-of-systems planning and synchronization.

While the scope of this document is limited to Army information systems, it is intended to be compatible and compliant with DoD and Joint-level interoperability requirements.

1.4 Objectives

The AIA has two (2) primary objectives:

- Provide the guidance and requirements, and identify the guidance and requirements documents, that govern the data-centric information sharing features of system architectures and ensure Army information systems meet the information sharing requirements of the Army; and
- Establish the requirements for assessing the compliance of systems to the guidance provided.

Additional objectives of this document include:

- Provide guidance that enables systems developers to meet DoD Directive 8320.02 [7] net-centric data sharing objectives of visibility, accessibility, understandability, trustworthiness, and interoperability;
- Present the priorities, principles, and business rules for information sharing in system architecture development;
- Identify the information sharing governance and guidance documentation¹ that (1) applies to Army information system design, development, deployment, and use; and (2) are derived from the AIA principles and priorities;
- Adopt and extend the architectural guidance provided in the DoD IEA into the Army system design and development;
- Identify and address gaps in Army system design and development guidance concerning data exchange; and
- Provide information that will guide IT investments.

1.5 Applicability

This document applies to new IT system solution design and development, and to upgrades/modifications to existing or legacy systems. This document does not require changes to legacy systems except as part of upgrades/modifications to those systems.

This document applies to some systems acquired by the Army insofar as it establishes system requirements for those systems. This document may not be applicable, for example, to many Commercial Off-The-Shelf (COTS) systems because the Army may not have been able to influence the design of the interoperability features of those systems. It would, however, be applicable to the integration of COTS product with other Army systems. This document will be applicable to those systems that are explicitly designed and acquired to meet specific Army performance/operational requirements; this document establishes the interoperability requirements for those systems. In the Army's acquisition process, the AIA guidance would be most applicable at Milestone B.

¹ "information sharing governance and guidance documentation" is a subset of Information Technology governance and guidance documentation as defined in Appendix B.2. The scope of information sharing governance and guidance is equivalent to the scope of the AIA.

1.6 Assumptions

This document is based on the following assumptions:

- The computing and network infrastructure needed to support the information architecture described by this document is defined by a combination of the LandWarNet 2020 End State Architecture [9], the COE Architecture [4], and by the activities described in the COE Implementation Plan [6].
- There will be continuing pressure to reduce system lifecycle costs while, at the same time, foster and improve interoperability.
- Programs and the development of the associated CEs will progress at different rates. Thus, the benefits derived from the implementation of AIA information sharing concepts, technologies, and standards, will emerge over time (i.e., there will be no “big bang” of interoperability benefits).

1.7 Audience/Stakeholders

The Army is the stakeholder and primary audience of the AIA. Specific organization and role-based stakeholders in the AIA are described below. The descriptions provide examples of reasons that the stakeholder is interested in the AIA and how they might use it. The descriptions do not assign responsibilities, but simply provide examples of stakeholder responsibilities.

Program Executive Offices (PEOs) are responsible for development, delivery, and deployment of individual Army systems. They ensure programs comply with the AIA and the COE Architecture as adopted by the Assistant Secretary of the Army (Acquisition, Logistics, and Technology) ASA(ALT) in the COE Implementation Plan. Their use of the AIA is to ensure the interoperability of deployed systems with other Army systems.

Program Managers (PMs) ensure that design guidelines provided in the AIA are applied to software system development within their programs.

System Architects and Developers follow AIA guidance in the design, development, and deployment of software systems. They are ultimately the point where “the rubber meets the road” when it comes to using/applying the AIA.

Data Stewards and Functional Data Managers (FDMs) provide a two-way communication pathway concerning AIA guidance. They bring awareness and explanation of the AIA back to the programs within which they are associated, and they also bring requirements up from the programs to the awareness of the broader Army. Data Stewards and FDMs are ultimately the primary source of input to the AIA.

Communities of Interest (COI) will use the AIA to analyze interoperability requirements and develop community-based interoperability solutions.

Army CIO/G-6 oversees information architecture standards and ensures that guidelines and tools are provided to implement systems that meet architecture requirements. CIO/G-6 is the organization responsible for the development/maintenance of the AIA, and it collaborates with ASA(ALT) to ensure standards and technical requirements are aligned to enable implementation.

Army Data Board (ADB), Army Data Council (ADC) and Chief Data Officer (CDO) oversee the Army data strategy, policies and practices and will use the AIA as the foundation for defining the data strategy and identifying data policies and practices.

Army Materiel Command (AMC) is responsible for Army material readiness, including technology, acquisition, materiel development, logistics, and sustainment. They would use the AIA in the development of logistics software systems that would, for example, oversee the aggregation and de-confliction of sensor data into a form more usable by analytic applications or send part change information to suppliers.

Assistant Secretary of the Army (Acquisition, Logistics, and Technology) (ASA(ALT)) is responsible for Army IT and is a main force behind the development of COE Architecture Implementation Plan to provide guidance on information and data exchange to PEOs and PMs. The ASA(ALT)'s primary concern is that the AIA provides guidance and measures to facilitate the interoperability of future software systems.

Army Network Enterprise Technology Command (NETCOM) provides information technology for all Army network communications. NETCOM plans, engineers, installs, integrates, protects and operates the Army's LandWarNet, enabling mission command through all phases of Joint, Interagency, Intergovernmental and Multinational operations. NETCOM builds and operates the technology foundation upon which the AIA guidance assumes and builds upon.

Office of Business Transformation (OBT) uses AIA guidance in business transformation planning (e.g., the OBT Business Transformation [9]).

Training and Doctrine Command (TRADOC) provides policy, guidance, and strategies to develop and sustain Army training and leader development systems. They establish the capability requirements that must be met by Army systems and, indirectly, the requirements for the AIA. They are also users of the AIA in the development of TRADOC software systems.

U.S. Forces Command (FORSCOM) is the ultimate user of the software systems developed based on AIA guidance and it is FORSCOM requirements that these systems must meet.

The AIA targets an Army audience, but the DoD and other DoD components, as well as joint forces, would also find the AIA valuable to meeting their information sharing objectives. Except for call-outs of Army-specific regulations and guidance, there is nothing Army-unique about the kind of guidance provided by the AIA, or the way that that guidance is organized. The AIA provides a data strategy that is applicable to any organization.

1.8 Army Information Enterprise Priorities

The DoD Information Enterprise is:

"...the DoD information resources, assets, and processes required to achieve the vision and perform the mission of the DoD CIO." ([2], p. 2)

Where the "DoD CIO's vision and mission are:

- Vision - Deliver agile and secure information capabilities to enhance combat power and decision making.
- Mission - Information is one of our Nation's greatest sources of power. Our first and greatest goal is to deliver that power to enable the achievement of mission success in all operations of the Department – warfighting, business, and intelligence." ([2], p. 2)

Similarly, the Army Information Enterprise is the information systems and processes that ensure that all Army personnel, from commanders to warfighters, get the right information at the right time to make the right decisions.

The DoD IEA [1] identifies five (5) priority areas for the transformation of current systems to net-centric operations that will support the Information Enterprise:

- Data and Services Deployment (DSD)
- Secured Availability (SA)
- Computing Infrastructure Readiness (CIR)
- Communications Readiness (CR)
- NetOps Agility (NOA)

In DoD IEA 1.2 [1], these priority areas played a prominent role in the structure and guidance provided by the DoD IEA. The AIA adopted the priority areas as a structuring mechanism for AIA guidance. The AIA provides data-centric information sharing guidance that focuses on the DSD and SA priority areas. The other priority areas are not directly about information sharing; rather, they are about computing technology and the performance of the technology and are outside the scope of this document.

In DoD IEA 2.0 [2] [3], the priority areas play a less prominent role. Instead, DoD IEA 2.0 provides high-level DoD Architecture Framework (DoDAF) architectural views (e.g., Capability Views) that describe the IEA; the DoDAF views are intended to be specialized via domain-specific Reference Architectures. The priority areas, principles, and business rules defined in the DoD IEA support elements of DoDAF architectural views. Table 7 in Appendix F of Volume II of DoD IEA 2.0 [3] provides a mapping of DoD IEA 1.2 priority areas to DoD IEA 2.0 Activities.

1.9 Document Overview

1.9.1 Overview of AIA Document Organization

The main body of the AIA provides the high-level guidance in the form of an End-State Information Sharing Framework and the definition of principles and business rules meant to guide system design and development efforts toward this end-state. More detailed guidance and compliance requirements are provided in the appendices. Figure 2 illustrates the structural overview of this document.

Section 2 explains how this document is intended to be used. There are two (2) principle uses: system development guidance and compliance assessment.

Section 3 presents the end-state system framework that describes the logical system functionality that compliance to the AIA guidance is intended to achieve.

Section 4 describes operational characteristics of Disconnected, Intermittent, and Limited (DIL) network environments and limited-capability CEs that may impact AIA guidance. Throughout the AIA, guidance is amended with changes necessary to accommodate the constraints of DIL operational characteristics.

Sections 5 through 9 present the principles and business rules that constitute the AIA guidance. The organization of the principles and business rules is described in detail in Section 1.9.2. These sections constitute the normative content of this document.

Section 10 describes the bodies and roles associated with Army data governance.

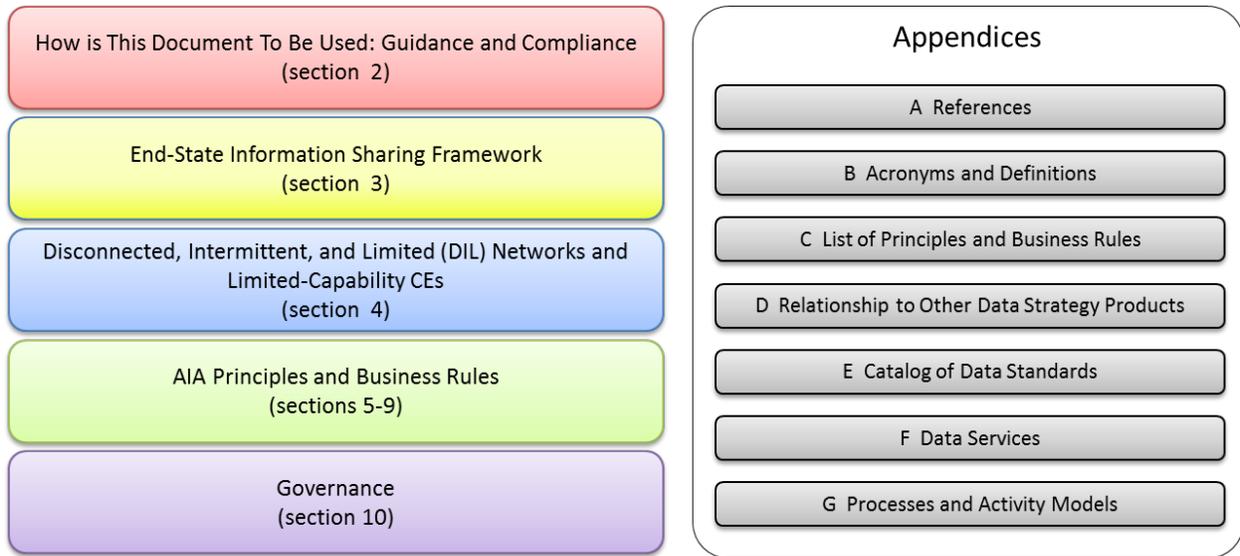


Figure 2: Document Organization Overview

The appendices provide information that supplements the main body of this document.

Appendix A provides the list of references cited throughout the document.

Appendix B provides the list of acronyms and definitions of terms used in the document.

Appendix C provides a simple list of the principles and business rules presented throughout the body of the document.

Appendix D describes the relationship of the AIA to other Army data strategy products.

Appendix E provides a catalog of data standards.

Appendix F describes data services, including data service standards.

Appendix G provides detailed guidance on subjects introduced in business rules. A business rule, in itself, cannot provide all the guidance necessary on a subject; therefore, many business rules reference additional material in Appendix G.

1.9.2 Overview of Principle and Business Rule Organization

Figure 3 provides an overview of the organization of the principles and business rules that comprise the guidance provided in the AIA and are presented in Sections 5 through 9.

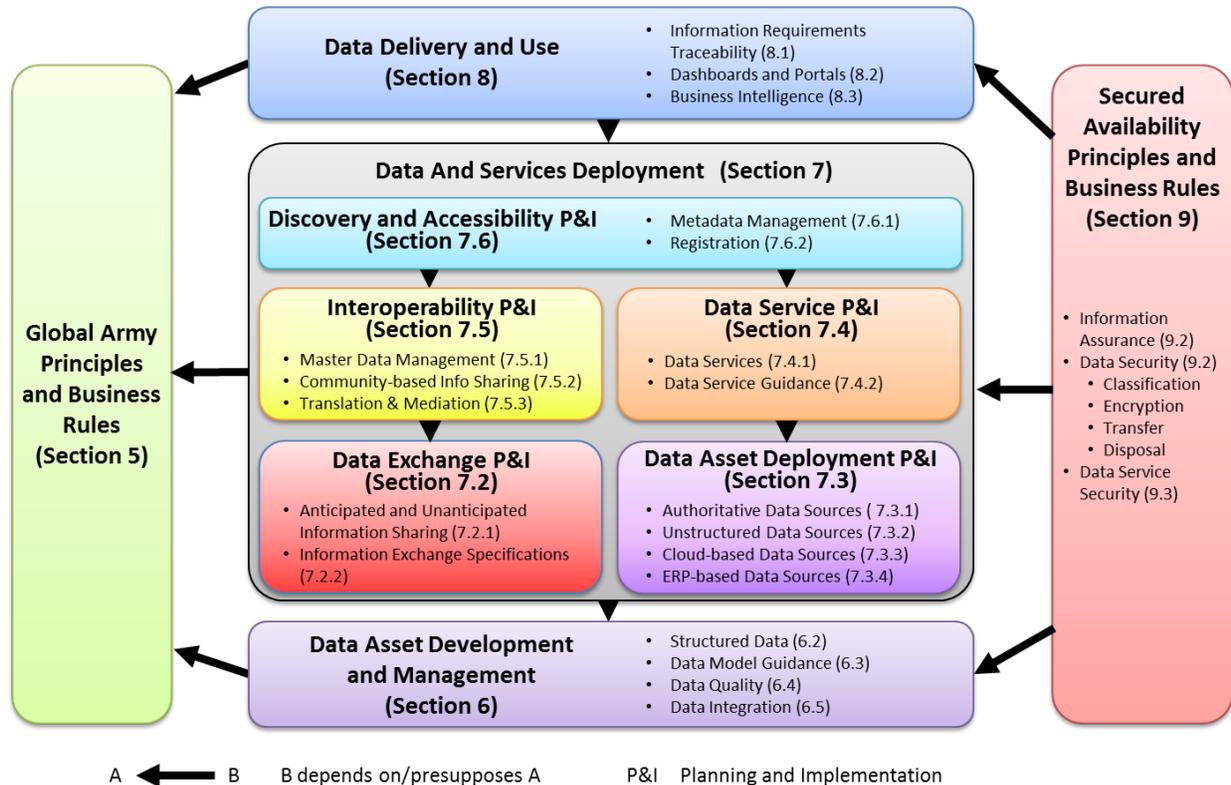


Figure 3: Organization of Principles and Business Rules

The principles and business rules are organized into five (5) primary subject areas and five (5) focused subject areas:

- Global Army principles and business rules are high-level data guidance that apply to all aspects of system design and development (see Section 5).
- Data Asset Development and Management principles and business rules are guidance that addresses data asset development directly as a precursor to data deployment, such as data modelling and data quality guidance (see Section 6).
- Data and Services Deployment principles and business are guidance on making data from data assets available to consumers throughout the Army (see Section 7); the guidance is provided in the following five (5) focused subject areas:
 - Data Exchange Planning and Implementation principles and business rules are guidance focused on anticipated versus unanticipated information sharing and Information Exchange Specifications (IES) (see Section 7.2);
 - Data Asset Deployment Planning and Implementation principles and business rules are guidance on making particular kinds of data assets, such as unstructured data assets, available to the Army as sources of data (see Section 7.3);
 - Data Service Planning and Implementation principles and business rules are guidance focused on data services development and implementation (see Section 7.4);

- Interoperability Planning and Implementation principles and business rules are guidance on enabling interoperability and information sharing within and between interoperability communities (see Section 7.5); and
- Discovery and Accessibility Planning and Implementation principles and business rules are guidance on metadata management and registering data and services with DoD registries to make them visible, discoverable, and accessible across the Army and DoD (see Section 7.6);
- Data Delivery and Use principles and business rules are guidance on the delivery and presentation of data to end-users (see Section 8).
- Secured Availability principles and business rules are guidance for ensuring that both data and data service access are secure and meet required DoD and Army security requirements (see Section 9).

The division of AIA principles and business rules into Global, Data and Services Deployment, and Secured Availability is intended to align with the way in which the DoD IEA organizes the DoD IEA principles and business rules. The Data Asset Development and Management and Data Delivery and Use areas are introduced in the AIA as siblings of the DoD IEA corresponding areas to provide better organized and more complete form of the AIA guidance. The Data Asset Development and Management area is internally focused on data assets themselves; Data Delivery and Use area is externally focused on the use and creation of data by Army end-users.

1.9.3 AIA Relationship to Other Documents

The relationship of the AIA document to other governance/guidance documentation is illustrated in Figure 4 and Figure 5.

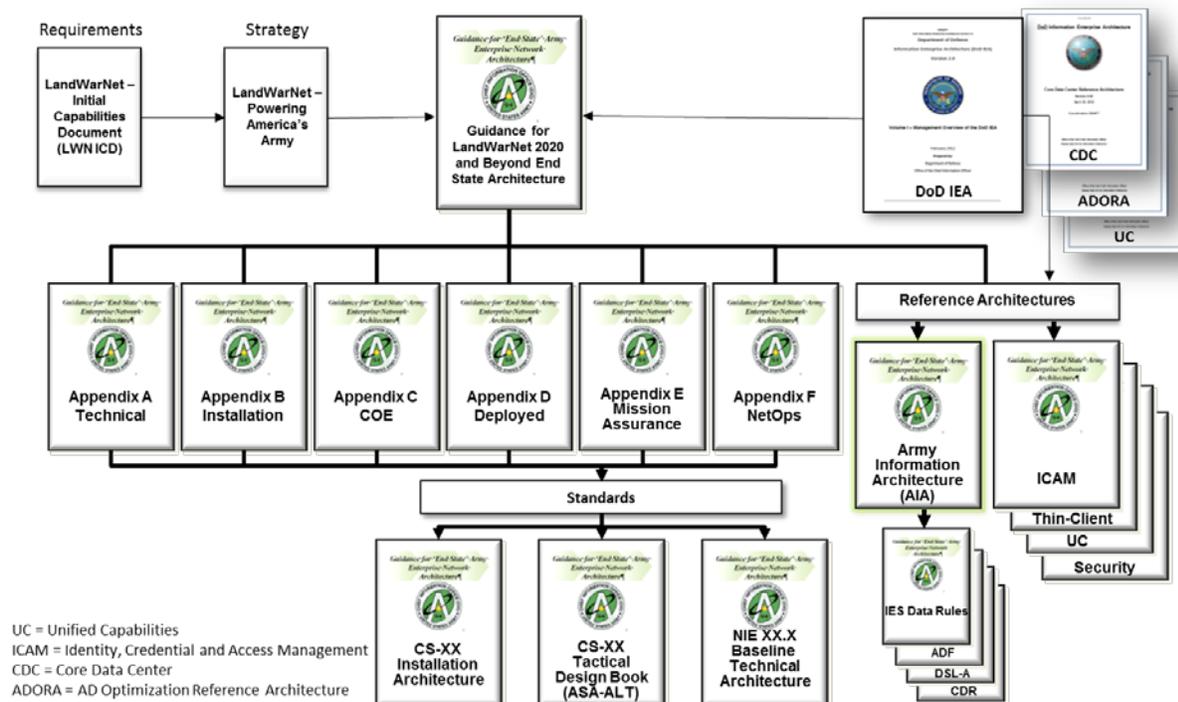


Figure 4: AIA Document Context

Figure 4 illustrates the positioning of the AIA within the context of Army IT architectural guidance documentation. The *Guidance for LWN 2020 and Beyond End State Architecture* [9] is the capstone guidance document that sets forth the vision for the Army's network architecture. It is supported by a number of appendices, one of which is the COE Architecture [4] (Appendix C), and it is complimented by a collection of reference architectures, one of which is the AIA. The DoD IEA is depicted in the upper right as one of several DoD/Joint level influences on the LandWarNet 2020 document and on the reference architectures. The other guidance documents depicted in Figure 4 are included for completeness of the context illustration, but it an explanation of each is not necessary for the purposes of this document.

Figure 5 illustrates the relationship of the AIA to guidance documents that have a direct relationship to the AIA. The documents and their relationship to the AIA are as follows:

- The **DoD IEA** [2] [3] establishes high-level architectural views, principles, and business rules for the definition of architectures that support transformation to net-centric operations. The AIA adopts the principles/business rules mechanism of the DoD IEA and aligns guidance with the DoD IEA Data and Services Deployment and Secured Availability priority areas (see Section 1.8).
- The **DoD Reference Architecture Description** [6] is a companion to the DoD IEA; it defines what a "reference architecture" is, describes the purpose and use of RAs, and identifies elements of an RA. See additional explanation of the relationship of the AIA to the DoD Reference Architecture Description below.
- Chairman of the Joint Chiefs of Staff (**CJCS**) **Instruction 6212.01F** [11] and **DoD Directive 8320.02** [7] provide the directives for how systems are designed and implemented to improve information-sharing capabilities among the systems. See additional explanation of the relationship of the AIA to DoD 8320.02 below.
- **Army Regulation (AR) 25-1** [12] and **DA Pamphlet 25-1-1** [13] specify Army policy for IT design and development.
- **Army Regulation (AR) 25-2** [14] and **(AR) 380-5** [15] specify Army policy for Information Assurance and security classification of information.
- The **Guidance for LandWarNet 2020 and Beyond End State Architecture** [9] (under development) is the master, umbrella guidance document for the future Army information system/network architecture. The architectural structure the AIA (see Section 3) aligns with the general architectural structure description of LandWarNet 2020 guidance.
- The **COE Architecture** [4] (Annex C of the *Guidance for LandWarNet 2020 and Beyond End State Architecture*) is a technology architecture document that provides guidance for the design of Computing Environments. The COE Architecture is a key element of the Army's overall system transformation plan. The COE Architecture is a technology architecture and the AIA information architecture complements that architectural perspective (see Section 1.3 and Figure 1.)
- The **AIA** defines priorities, principles, business rules, and technology requirements related to information sharing for Army system architecture design and specification.
- The AIA cites the following technical guidance documentation that is input to and leveraged in the design of solution architectures and specifications:
 - **Army Data Framework (ADF)** [14];
 - **Data Strategy Reference Architecture (DSRA)** [14];
 - **Data Services Layer – Army (DSL-A)** [18];
 - **Content Discovery and Retrieval (CDR)** [19] [20];
 - **Namespace Management for the Army Enterprise** [21] [22]; and
 - **Rules for Cross-Cutting Capability (CCC) Information Exchange Specifications (IES) in Interface Specifications ("IES Data Rules")** [23].

- The **COE Implementation Plan** [6] provides programmatic guidance for the phase implementation of the COE Architecture, including additional architecture drill-down guidance and implementation and execution plans.
- Information System Architectures (e.g., DoD Architectural Framework (DoDAF) views [19]) specify and organize system specifications.
- System specifications govern how the system is built.

The COE Implementation Plan, system architectures and specifications, and a Work Breakdown Structure (WBS)/schedule govern the System Development Effort that produces the real-world system.

The bolded document icons shown in Figure 5 describe the Army data architecture.

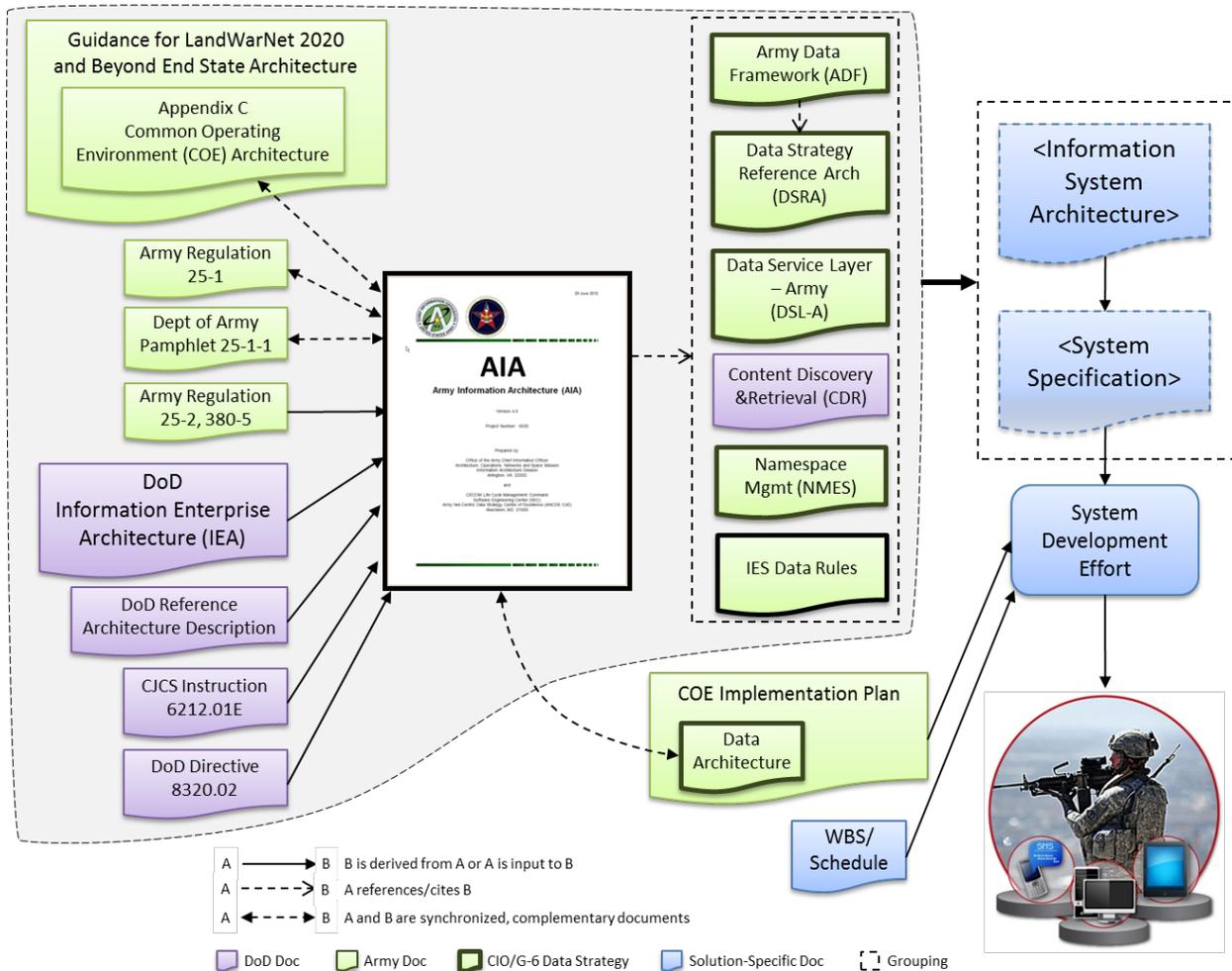


Figure 5: Relationship of AIA to Other Document Products

DoD Directive 8320.02 [7] establishes five (5) net-centric data sharing objectives as policy:

- Visibility
- Accessibility
- Understandability
- Trusted
- Interoperability

The AIA directly supports these objectives. For each objective, Table 1 identifies the sections of the AIA that provide guidance directly supporting that objective.

Table 1: Relationship of AIA to DoD Net-Centric Data Sharing Objectives

Objective	Relevant Section of the AIA
Visibility	Section 7.6.1 Metadata Management Section 7.6.2 Registration
Accessibility	Section 7.4.1 Data Services Section 7.4.2 Data Service Guidance
Understandability	Section 6.3 Data Model Guidance Section 6.5 Data Integration Section 8.2 Information Requirements Traceability
Trusted	Section 6.4 Data Quality Section 7.3.1 Authoritative Data Sources (ADS) Section 9 Secured Availability (SA)
Interoperability	Section 7.2.2 Information Exchange Specifications Section 7.5.1 Master Data Management Section 7.5.2 Community-based Information Sharing Section 7.5.3 Translation and Mediation

The DoD Reference Architecture Description [6] identifies five (5) elements of an RA, which are presented in the first column of Table 2. The second column of Table 2 identifies the features of the AIA that matches the RA element.

Table 2: Relationship of AIA to DoD Reference Architecture Elements

RA Element	Corresponding Feature of AIA
Strategic Objective	The strategic objective of the AIA as an RA is presented in Section 1.
Principles	AIA principles
Technical Positions	AIA business rules
Patterns/Templates	The AIA identifies and defines patterns as appropriate, e.g., the Data Service development process presented in G.3.1.
Vocabulary	The AIA defines terminology throughout the document; a complete list of defined terms is presented in Appendix B.2.

2. How Is This Document To Be Used?

2.1 Guidance

This document provides guidance on the design of information system architectures and specifications in the form of principles, business rules, and references to guidance presented in complimentary documentation. The guidance focuses on the information sharing and data exchange features of architectures and specifications and assumes that guidance for other features of the system under design is provided by other guidance documentation.

A principle is a generalized statement of position that is accepted as true or valid, and often reflects values, beliefs, or convictions on the “right” or “best” way to do or achieve a result or fulfill a mission. Principles guide decision-making and actions; a principle is not an end-state objective. This description is compatible with the definition of “principle” as provided in the DoD Reference Architecture Description [6].

Business rules are by-products of principles. They are recommendations, requirements, guidelines, directives, stipulations, or imperatives that assert what shall/should be done to meet or implement the principle. A business rule may be an end-state objective. There may be many different business rules that meet/implement a principle, but those identified in this document will be the ones adopted and used by Army system development teams. This description is compatible with the definition of “technical position” as provided in the DoD Reference Architecture Description [6].

The appendices provide additional guidance that complements the main content of this document. Business rules may cite material in an appendix or external documentation (e.g., standards) as additional guidance.

2.2 Compliance

This document provides a basis for assessing and measuring compliance with Army IT architectural guidance pertaining to information sharing and data exchange. Compliance with the AIA requires an understanding of and adherence to the business rules contained herein. If material in an appendix or external documentation is cited in a business rule, then compliance with the requirements of that material is also required. Compliance with DoD IEA [2] [3] is assumed as a precursor to AIA compliance. In addition, Army PMs and Programs of Records (PoRs) are expected to comply with the Army’s COE Architecture requirements [4] and higher level Army and DoD directives and regulations.

The business rules presented throughout the AIA document are the basis for AIA compliance assessment. Compliance assessment is the evaluation of an item (e.g., organization, product, system, data asset, or data service) for adherence to the AIA business rules. For Army management, compliance assessment provides a measure of “how net-centric” Army systems are; for the assessee, compliance assessment provides a checklist of Army guidance on what it means to be “net-centric.”

A compliance matrix has been created and used to assess system compliance against AIA business rules. The matrix and process of using it is an evolving capability; the current state of this capability is documented by AIA Compliance Matrix and Assessee Briefing Deck [25].

Waivers or exceptions to AIA compliance shall follow the waiver/exception process specified in the COE Architecture [4].

Compliance assessment may be performed at any point in the system lifecycle. Early in the system lifecycle, compliance assessment will identify design guidance that needs to be accounted for as the system is designed and built. Later in the system lifecycle, compliance assessment will identify opportunities for system evolution and improvement.

NOTE: The DoD IEA [2] [3], COE Architecture [4], and AIA document are the primary source of compliance requirements that must be met by Army systems in order to meet Army and DoD net-centricity and information sharing objectives. However, they are not an exhaustive set of specifications to which Army systems must comply. Other compliance requirements will be generated in/by other DoD and Army governance activities. For example, the cross-cutting Army Geospatial Enterprise must align with the information aspects of the National System for Geospatial-Intelligence (NSG) Enterprise Architecture as well as the AIA.

3. End-State Information Sharing Framework

The vision of the AIA is the definition of the bottom-to-top fabric of how the generating and operating forces use data to provide, share, and use accurate and actionable information to enhance mission effectiveness. The end-state vision is a flexible, adaptable, and robust system that:

- (1) delivers data from any of the Army's diverse data assets to any Army consumer (human user or application); and
- (2) is designed and implemented based on Service-Oriented Architecture (SOA) principles.

SOA system design principles are well-suited to achieving DoD objectives for information sharing. For example, Loose Coupling and Reusability [25] enable and support flexible and adaptable system behavior.

The vision is represented by the End-State Information Sharing Framework illustrated in Figure 6 and Figure 7. Figure 6 illustrates a simple form of the framework, presenting the end-state as a three-layered SOA-based architecture. The bottom layer, which serves as the foundation of the architecture, is the data layer; the data layer is comprised of diverse data assets that house, maintain, and supply data to Army processes. The middle layer is the service layer; the service layer is comprised of data services that (1) provide access to data layer resources and (2) are available to consumers throughout the Army. The top layer is the application layer; the application layer provides the business and mission functionality that is available to users in the Army.

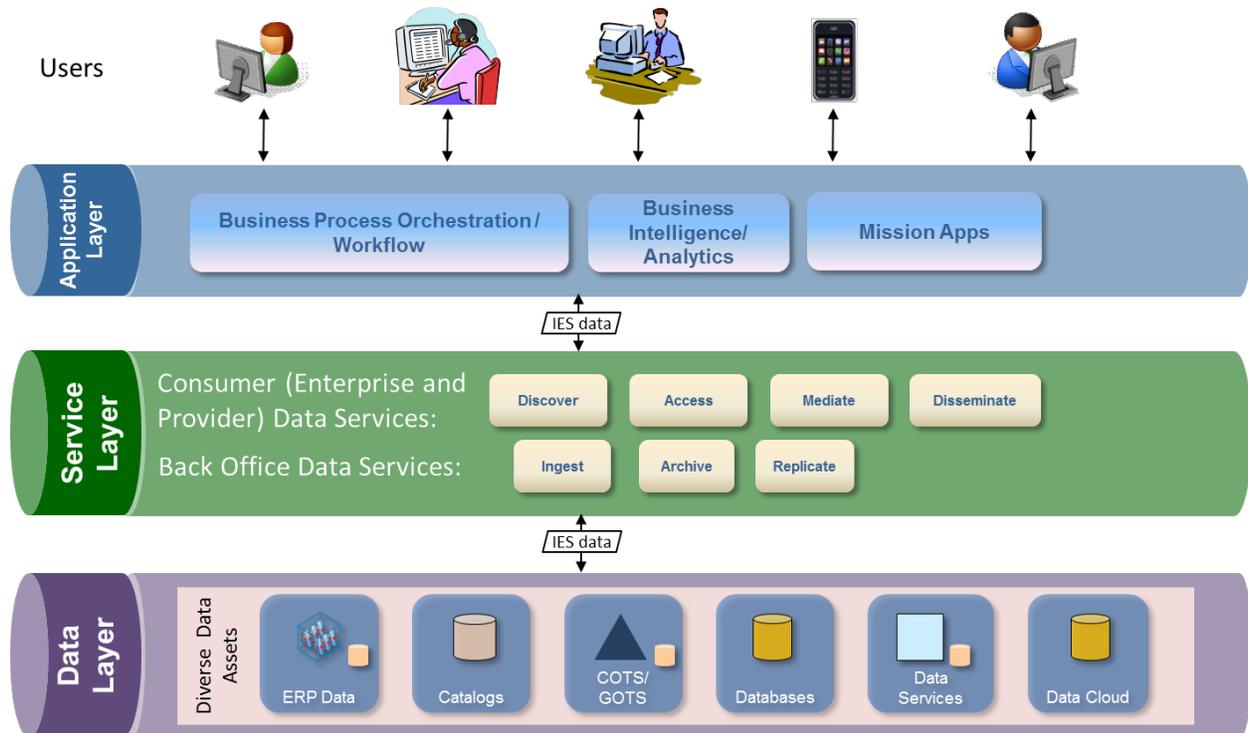


Figure 6: End-State Information Sharing Framework (Simple Form)

Figure 7 expands on the illustration presented in Figure 6 and presents a logical view of the end-state information sharing functionality. The three (3) layers of Figure 6 are also illustrated in Figure 7 with shadowed versions of the layer boundaries. The diverse data assets illustrated

in the data layer of Figure 7 are just a few of the many different possibilities showing how data may exist and be managed within the Army. What is common to all data assets is a service interface that provides simple Search-Create-Read-Update-Delete (SCRUD) functionality and access to the data. The Search and Read functionality is assumed for each data asset as a minimum; the Create, Update, and Delete functionality is optionally available as appropriate to authorized users. The SCRUD service interfaces “plug into” the network (represented by the thick dark line in the illustration) and are available to services and other consumers at all levels of the architecture.

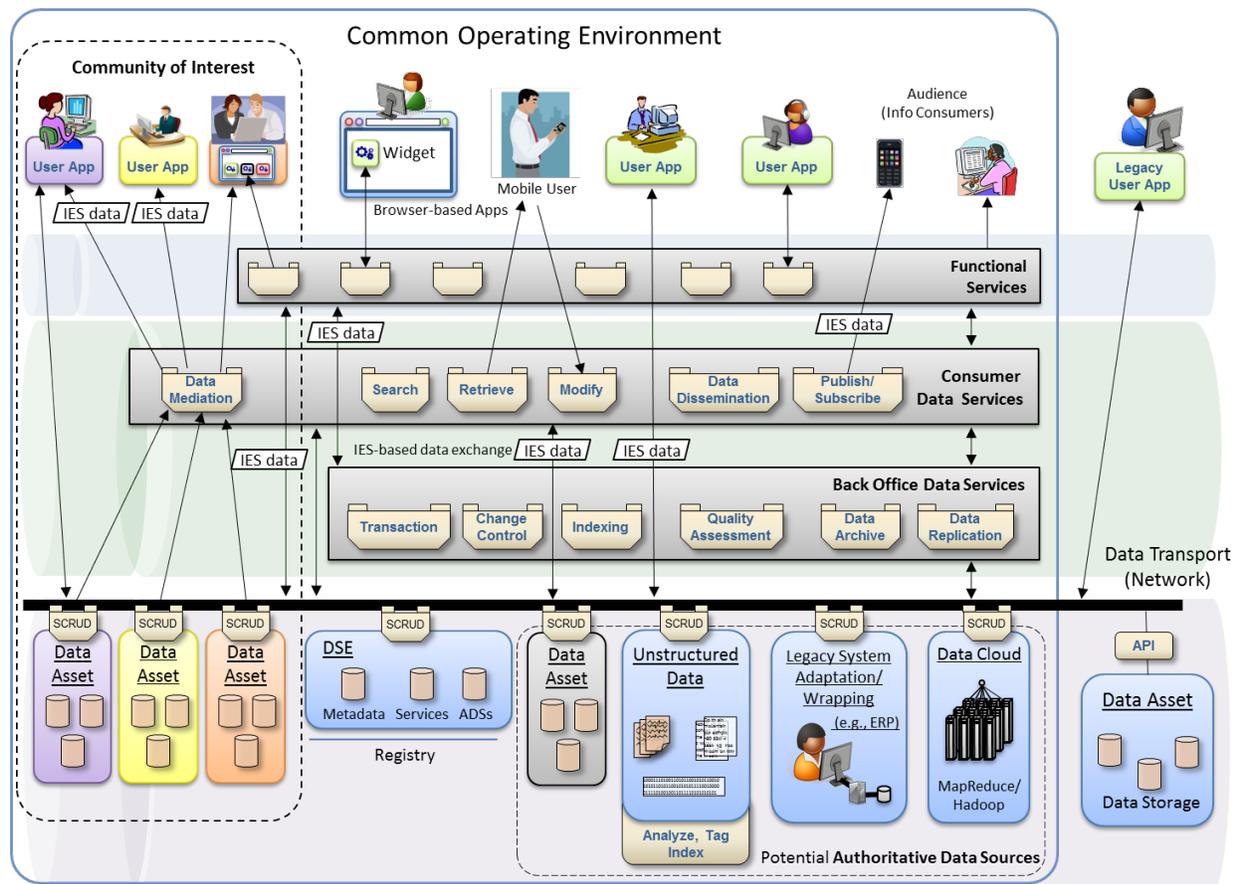


Figure 7: Army Information Architecture (AIA) End-State Information Sharing Framework

The service layer illustrates more example services. Services are grouped into several categories based on the functionality offered by each service. Consumer data services are visible and usable by consumer applications and end-users. Back office data services are data management services not used by most consumers; they provide critical infrastructure capabilities to manage and secure the data assets. Consumer data services can be further categorized as enterprise data services and provider data services. Enterprise data services are implemented and centrally managed (e.g., by a service portfolio manager) at the Army level, such as a VMF<->XML conversion service. Provider data services are services that are developed and deployed to meet Army functional needs, but are not centrally managed.

The application layer consists of functional services and applications. These capabilities have a smaller audience and provide specific capability that is often unique to the user functional domain. Service orchestration, in which multiple services are combined to provide a complex end-user capability or business service, also takes place in the application layer.

Users are not just consumers of data, but may be creators of data as well. For example, situational awareness data gathered by forward units on mobile devices may be pushed into the system, captured into appropriate data assets, and disseminated out to affected parties.

Other features illustrated in the diagram include the following:

- Interactions between services (illustrated by the arrows) are realized by the secure exchange of data (e.g., the payloads of messages exchanged between services, or files transferred between systems). The exchanged data should be based on IESs.
- A group of end users, their systems, and their data may form an informal interoperability community or a formal COI to facilitate and manage interoperability within the group and between the group and other external communities.
- The designation of a data asset as an Authoritative Data Source (ADS) is a role played by the data asset. Any of the data asset types illustrated may play the role of an ADS.
- Services may be connected to widgets (i.e., User Interface (UI) components that provide specific functionality or information) that users add to their browsers, enabling them to create a mashup (i.e., combination of diverse components) of information from different service-based data assets.
- Interoperability with systems and data assets external to the COE architecture is possible through customized access to Application Program Interfaces (APIs) of the external systems.

The guidance provided in the AIA is intended to direct systems development in such a way as to facilitate, enable, and promote the realization of this framework.

The organization of the AIA guidance as illustrated in the three center layers of Figure 3 correspond to the three (3) layers of the End-State Information Sharing Framework illustrated in Figure 7. The correspondence makes it easier to see how the AIA guidance contributes to the realization of the end-state vision.

4. Disconnected, Intermittent, and Limited (DIL) Networks and Limited Capability CEs

The End-State Information Sharing Framework presented above and the information sharing guidance provided in the AIA assumes a “most capable” network environment in which bandwidth, connectivity and computing capability are not an issue. Many parts of the Army operate in physical locations and environments that are far from this ideal, such as warfighting units operating in theater (often called the “Tactical Edge”). The networks operating in these environments are called DIL networks to describe their limited connectivity characteristics. The Army also employs platforms and devices that range in computing power from the upper-bound capability of a data center to the limited capability of hand-held devices and sensors. The AIA guidance must therefore be tempered, modified, or possibly suspended, to suit the conditions of DIL networks and the limited capabilities of non-data center platforms/devices where such guidance may not be applicable or may be detrimental..

The COE Architecture [4] defines six (6) CEs with various levels of capabilities:

- Data Center / Cloud / Generating Force;
- Command Post;
- Mounted;
- Mobile/Hand Held;
- Sensor; and
- Real Time / Safety Critical / Embedded.

The Data Center CE is assumed to be fully capable and not limited in any way. All AIA guidance would apply to Data Center CEs, relationships between them, and networked applications/sites in Continental United States (CONUS) and coalition regions.

Non-data center CEs are mobile and may move in and out of DIL network environments. The DIL connectivity and non-data center CE computing capability are two (2) independent factors to consider with respect to the applicability of AIA guidance. Non-data centers will be collectively referred to as *Limited CEs* (LCE) throughout the remainder of this document.

The Command Post CE is a transportable environment and is a very capable and generally reliable network with moderate to high bandwidth located in command posts or improved building environments. Command Post CE connectivity may be partially limited due to location and reliance on satellite connectivity.

Mounted, Mobile/Hand Held, Sensor, and Real Time CEs often operate in DIL network environments. The computing capability of these CEs is much more limited than the Data Center and Command Post CEs.

The simplest example of how DIL environment characteristics would impact AIA information sharing and data exchange guidance is the use of Extensible Markup Language (XML) for data exchange. From a bit-count/file size perspective, XML messages are extremely verbose and are far too large to practically exchange over networks where connectivity or bandwidth is limited. Therefore, data exchange with LCEs in DIL environments, particularly handhelds, should use a compressed binary format if compatible with compression/uncompression latency.

As guidance (e.g., principles and business rules) is presented through the document, amendments, refinements, replacements, or waivers of that guidance for DIL environments and LCEs will be presented using the following presentation convention:

DIL/LCE	Amendments, refinements, replacements, waivers, or other commentary regarding the applicability of a business rule in or with respect to a DIL environment or an LCE will be provided in a boxed paragraph like this.
----------------	---

If there is no material provided about DIL CEs in conjunction with the guidance provided, it shall be assumed that the guidance applies to all CEs.

5. Global Principles and Business Rules

Global principles and business rules that apply to information sharing within the Army are presented in the following subsections. These principles and business rules are overarching concepts that span and provide a basis for the principles and business rules in subsequent sections.

5.1 Data and Information

The terms *data* and *information* are often conflated or used synonymously. For the purpose of AIA, a distinction is made between them in order to separate solutions from business requirements.

Information is (a) derived from data when data is interpreted within the context of a mission process, or (b) encoded in data when data is created within the context of a mission process; information is, thus, context sensitive. Information is intangible; data is tangible and physically observable. Information flows between enterprise business processes; it is what is communicated between collaborating agents, and is the basis of decision-making. Data is the physical mechanism that conveys information and is the physical signal exchanged between sender and receiver; data is the physical item exchanged between information systems.

Knowledge, simply, is what people know – knowledge exists in a person’s head. Information is interpreted from data within the context of a person’s knowledge, which is why two people may read the same sentence yet derive different information from it.

Principle GA-01: Data is an Enterprise Asset. Information is Enterprise Currency. Knowledge is an Enterprise Resource.

Data is an asset because it is something of value that is tracked and managed. Information is currency because it is something of value that is exchanged between parties for the mutual benefit of both parties. Knowledge is a resource because it is something of value that is drawn upon and used to achieve an end; knowledge is not an asset because it cannot objectively be tracked and managed.

Principle GA-02: Data is a physical representation of information but is not the same thing as information.

The distinction between information and data is reflected in the definition/explanation of DoDAF [19] architectural views. In the DoDAF Capability Views (CVs) and Operational Views (OVs), information is communicated between capabilities or processes along need-lines. In the DoDAF System Views (SVs) and Service Views (SvcVs), data is exchanged between systems, software applications or services via interfaces. The distinction is also reflected in the Army *Information Security Program*, AR380-5 [15], which states that information is classified at a certain security classification level, e.g., “the range of missile XYZ is classified as Secret,” and that the various forms or expressions of that information retain that classification, whether spoken, written in a different language, or put into a database.

These principles provide the basis for differentiating between several important functions that comprise information sharing. They separate pure data management (e.g., backup) from providing information to end-users (e.g., consumer service access). They also provide the basis for understanding the need for data translation (see Section 7.5.3) because the same information may be represented by data that is structured differently.

5.2 Data Exchange and Information Sharing

Following the distinctions above, *data exchange* refers to the rote, mechanical, physical transference of data without consideration of meaning or intent. *Information sharing*, on the other hand, is the exchange of data with the purpose of conveying, sharing, or communicating information pertinent to some purpose or process. Information sharing deals with meeting business needs. Data exchange deals with moving data between software applications. Because information sharing cannot happen without “data” exchange (where “data” is interpreted broadly to include verbal and written communication in addition to digital data), the two terms are often used synonymously.

Effective information sharing throughout the Army and across the DoD is the primary goal of the guidance provided in the AIA and the primary goal of the DoD information sharing objectives.

Principle GA-03: Effective decision-making and effective process execution in the Army requires effective Information Sharing.

Principle GA-04: Information creators and managers have a responsibility and obligation to make their data visible and accessible to consumers throughout the Army.

Business Rule GA-04a: Information creators and managers shall have a plan and schedule (i.e., implementation plan) for making their data available to the Army (if not already available).

DIL/LCE	For users/devices in LCEs, the plan for making data available to the Army (or the need for such a plan) is dependent upon the operational situation of the users/devices and requires an analysis of the situations and the data to be uploaded. In general, the plan for LCEs should be to upload all new data to more capable CEs during periods of favorable connectivity.
----------------	---

The data assets in the bottom layer of Figure 7 illustrate data that is available to the Army through a SCRUD interface accessible over Army networks.

Making data available to the Army is distinct from ensuring that the data is accessible by consumers throughout the Army. Making data available entails a willingness and ability to share data with consumers; making data accessible entails making data available through a standardized mechanism that is known/expected by consumers.

Principle GA-05: The information that drives decision-making and Army processes is available to authorized consumers regardless of their location or the time of their request.

Business Rule GA-05a: Data should be accessible by authorized consumers across the Army within the security restrictions on the data.

The users across the top of Figure 7 illustrate some of the kinds of consumers and ways in which the consumers may access data services.

DIL/LCE	Consumers using devices in LCEs should be able to access data like any other consumer, with the only difference being response time due to the limitations of a DIL environment.
----------------	--

5.3 Information Sharing Governance and Guidance

Consistent adherence to and compliance with governance (mandatory) and guidance (recommended) documentation, tools, and other resources will direct and lead to the convergence of Army systems toward meeting the Army's end-state information sharing framework.

The ADB is the ultimate authority for information sharing governance and guidance documentation and adjudication of governance/guidance issues. See Section 10 for an explanation of the ADB and the governance function.

Either data stewards or bodies designated by the ADB will be responsible for information sharing governance and guidance documentation, including development, coordination, testing, promotion, sustainment, and compliance assessment. The development of governance/guidance shall align or comply with any applicable higher-level guidance; for example, data quality guidance developed by data stewards shall align or comply with any data quality guidance adopted/approved by the ADB, the Army, and/or the DoD.

There is a pattern of principles and business rules that recur throughout the AIA. The pattern, of which GA-06 is the first example, is of the form:

Principle: Governance/guidance documentation helps with XYZ.

Business Rule 1: Data stewards or ADB designee should develop governance/guidance documentation on XYZ.

Business Rule 2: Architects and developers should follow governance/guidance documentation on XYZ.

Each instance of the pattern save the first (GA-06) is a specialization of a previous instance of the pattern; most of them are specializations of GA-06. It is beyond the scope of the AIA, however, to devise a taxonomy of governance/guidance documentation.

Principle GA-06: Compliance with Army governance and guidance documentation will enable, facilitate, and promote effective information sharing among Army information systems and meet DoD information sharing objectives.

Business Rule GA-06a: Data stewards or an ADB designee shall develop, maintain, and promote data, service, and architecture governance and guidance documentation and shall assess compliance to the documentation.

Appendix C provides a description of existing governance/guidance documentation that is related to the AIA and how it is related.

Business Rule GA-06b: Architects and developers shall ensure that systems comply with the following data governance and architectural guidance documentation, as applicable:

- Army Knowledge Management and Information Technology, AR 25-1 [12];
- The COE Architecture [4];
- The Army Data Framework (ADF) [14];
- The Data Strategy Reference Architecture [14];
- Content Discovery & Retrieval (CDR) [19]; and
- Data Services Layer - Army [18].

Business Rule GA-06c: Data stewards or an ADB designee should develop/acquire, test, and promote tools and resources that support adherence to or compliance with data, service, and architecture governance and guidance documentation.

Business Rule GA-06d: Architects and developers shall adopt, implement, or use standards and governance/guidance documentation in the preferential order presented in Table 3 (adapted from [57]).

Table 3: Standards Adoption Priority

Priority	Level/Scope	Example Classification/Source
1	International	International Standards Organization (ISO); International Electrotechnical Commission (IEC); International Telecommunications Union (ITU)
2	National	American National Standards Institute (ANSI)
3	Professional Society; Technology Consortia; Industry Association	Institute of Electrical and Electronics Engineers (IEEE); Internet Engineering Task Force (IETF); World Wide Web Consortium (W3C); Organization for the Advancement of Structured Information Standards (OASIS); Government Electronics & Information Technology Association (GEIA)
4	Government/Federal	Federal Information Processing Standards (FIPS)
5	Military/DoD	Military Standards (MIL-STDS); Standardization Agreements (STANAGS); DoD Directives, Instructions, Manuals, and Guides; DoD Information Technology Standards and Profile Registry (DISR)
6	Military/Army	Regulations, Directives;

DIL/LCE	<p>Business Rule GA-06e: Army governance and guidance shall take into account constraints of LCEs and DIL environments.</p> <p>Business Rule GA-06f: If the governance and guidance documentation does not explicitly address LCEs or DIL environments, LCEs should comply with the documentation cited in GA-06b as applicable to/within the constraints of the LCE and potential DIL environments.</p>
----------------	--

The effectiveness of Army IT governance/guidance documentation, including the AIA, must be assessed to ascertain the value of that documentation. If the governance/guidance documentation is not materially improving the ease with which systems are developed, reducing the cost of system development, and improving the quality of those systems, then the documentation is of no value.

Principle GA-07: The effectiveness of Army governance documentation can be measured (in part) by the cost savings that results from adopting the guidance/solutions.

Business Rule GA-07a: PoRs/PMs/Data Stewards should separately track costs of development, deployment, and sustainment of enterprise data, data services, and COI activities, to measure, manage, and improve efficiency and effectiveness of AIA.

For example, PMs should be able to report on the costs of sustaining enterprise data assets, of developing data services, and of engaging in COIs.

6. Data Asset Development and Management (DADM)

6.1 Foundation for Data Exchange and Information Sharing

All complex structures must be built on a solid foundation that can support both the construction and operation of the structure eventually built on it. For information sharing in the Army, the foundation of the AIA is the data assets that support Army systems. Effective information sharing starts with well-designed, well-managed, and well-maintained data assets.

DoDD 8320.02 [7] defines a “data asset” as:

“Any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a web site that returns data in response to specific queries (e.g., www.weather.com) would be a data asset. A human, system, or application may create a data asset.” [7]

The term “data asset” as used in this document complies with this definition of “data asset.” Figure 8 illustrates examples of data assets within a simple organizing structure.

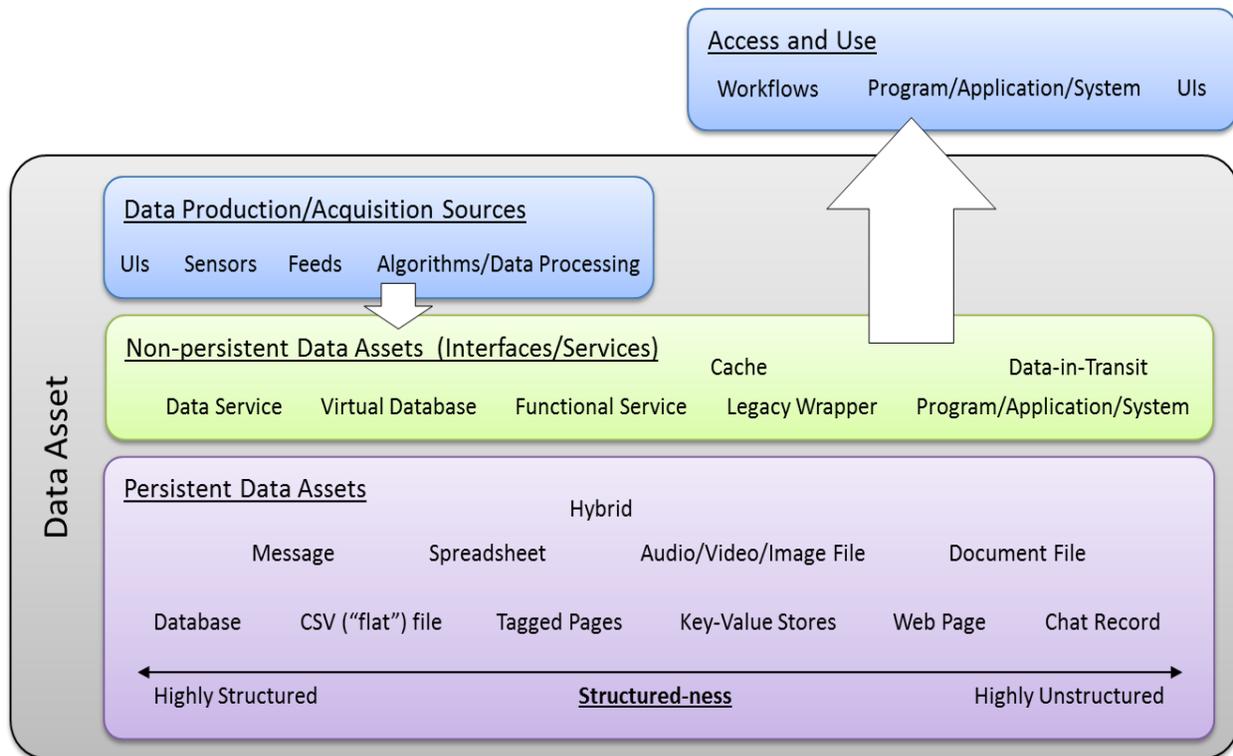


Figure 8: Examples of Data Assets

Before a data asset is deployed and made available to consumers, via a file or a data service interface, there are several “inward looking” items that should be considered and addressed. These items pertain to the data asset itself and include the following:

- **Physical Structuring of Data:** The choice (whether intentional or inherited) of how data is physically structured directly affects the ability of software and people to extract or encode information in the data. Structured data makes it easier for software applications to interpret and store information; unstructured data typically requires human interpretation to extract information.
- **Use of Data Model Design Resources.** Data models are the design specifications for data. The quality (or lack thereof) of data models directly affects the quality and usability of data. The Army makes Data Model Guidance (e.g., guidelines, standard vocabularies, reusable schema components) available to data asset developers/maintainers to guide data model design. This guidance includes data dictionaries, concept taxonomies, and data model fragments. The objective of the development and use of these data model design resources is to promote and engender common practices and perspectives within the data modelling community. See Section 6.3 for principles and business rules associated with data model design resources.
- **Data Quality:** Bad data means bad information, and bad information means bad decisions. Data quality is critical to enterprise effectiveness and is affected by many factors, from the design of the data models to user input. Data Quality Management (DQM) is essential to effective information sharing.
- **Data Integration:** Data coming “into” a data asset or system from external sources contains information that needs to be integrated into/with the information already contained in the data asset/system. The incoming data is often structured differently (i.e., using a different data model) than the local data asset and must be transformed to the format of the local data asset before integration. Data integration is the process of combining incoming data from one or more external data assets with the data in a local data asset and producing a single, unified, consistent, and cohesive data set that accurately and correctly represents the combined information content of the incoming data and the local data.
- **Master Data Management:** Master Data Management (MDM) is the set of enterprise functions for tracking, controlling, and maintaining data about persistent, non-transactional objects important to the enterprise. The scope of these functions is all of the data assets of the enterprise.

The following sections describe and discuss these items and present the principles and business rules associated with them.

6.2 Physical Structuring of Data

The physical structure of data takes many forms throughout the Army, from relational databases to XML documents to wikis and social media sites. All data provides information that may be valuable to the Army; obtaining and sharing this information depends on strategies and techniques for dealing with the particular way that the data is structured.

For the purposes of the AIA, data can be described as falling into two (2) broad categories: structured and unstructured. Structured data is data that can be governed by a schema and can be validated against that schema; structured data typically represents “fine-grained semantics” where a data element name indicates a localized meaning of a data value at a level that can be used by software applications.

Unstructured data is text, video, audio, or imagery that is intended for interpretation by human agents. The text is typically natural language expressions that would be used in email, social networking sites, news sites, etc.

Between structured data and unstructured data is a continuum of “structured-ness”; data within this continuum is called *semi-structured data*. Semi-structured data has structural characteristics that enable software applications to extract some information from the data, though the full extent of information available from the data requires human interpretation. Examples of structural characteristics that add structure to unstructured data include tagging unstructured data with metadata or the positioning of unstructured data at certain locations in the dataset.

Principle DADM-01: Information of value to the Army is represented by structured, semi-structured, and unstructured data.

Data that is stored persistently on physical media is often called “data-at-rest.” Data-at-rest includes databases or files on internal/external magnetic media, optical media, Universal Serial Bus (USB) drives, or solid-state drives. Data that is transmitted over a network (physical or wireless) is often called “data-in-transit.” Data-in-transit includes messages en route from source to destination, signals, and streams. Some data, like messages, can transition from in-transit to at-rest many times while moving from source to destination, but unless the “at-rest” duration is significant, is usually considered “in-transit” throughout.

Structured data is the primary mechanism by which software applications process information. Human programmers write software to “interpret” the data and execute the appropriate action based on the information interpreted from the data. The physical, persistent storage of structured data is the primary mechanism by which software applications store information for subsequent reuse. The physical exchange of structured data is the primary mechanism by which software applications share information.

Managing structured data is critical to information sharing and interoperability within the Army.

Structured data exists within a bounded container such as a database, a file, or a message. “Bounded” means that the container has finite, physical boundaries such that it can be determined whether some piece of data is either “inside” or “outside” the container. A structured data asset is a container that meets the following conditions:

- is governed by a single schema (i.e., physical data model), or set of integrated schemas, which entails that all data type names for the data elements are unique;
- the boundaries of the container define a managed identifier space within which all data element identifiers (e.g., relational data keys, record IDs, or XML “ID” attributes) are unique; and
- has a single unique, holistic identifier (e.g., path/file name, database ID, message ID).

Since “data asset” refers to anything from which data is available, a data asset may be a data service interface, a virtual database, or federated database. Non-persistent data assets such as these can be considered as meeting the above conditions if the data assets “behind” them meet the conditions and the conditions are managed/integrated from those data assets to the non-persistent data asset.

Principle DADM-02: Effective information sharing is based on clear, unambiguous, and consistent management of structured data. Physical data models (aka schemas) are a necessary mechanism for managing the format and semantics of (i.e., the information conveyed by) structured data.

Business Rule DADM-02a: A schema (i.e., physical data model) shall be developed and maintained for each structured data asset (e.g., database, data service interface, or message format). In DoDAF architectures, the schema would be a DIV-3 Physical Data Model.

Business Rule DADM-02b: An inventory (list) of the data assets within the scope of responsibility of a program or system shall be developed and maintained.

See Section 6.3 for guidance on data model development.

6.3 Data Model Guidance

A significant contributor to the “stove-piping” of systems – and resulting inability to effectively share information – is the diversity of understanding, perspectives, experience, technologies, and approaches that are used in the development of Army information systems. The Army is addressing this problem, in part, with the development and publication of the COE and the AIA. Given the critical roles of data models in the development of interoperable applications, additional development guidance is needed that focuses directly on the data models used for information sharing. Data Model Guidance is the collection of procedures, methods, best practices, recommendations, and subject matter expertise that support data model design, development, and implementation. Data model guidance includes, but is not limited to:

- Standardized, reusable schema components for ubiquitous concepts (e.g., person, location, time) that can be incorporated into data models under development;
- Vocabularies, Taxonomies, Data Dictionaries, and Glossaries that establish common definitions of common terms and their hierarchical relationships;
- Ontologies to define material domains with semantic precision;
- Guidelines:
 - Naming conventions;
 - Modelling patterns, paradigms, styles;
 - Design and analysis principles and practices;
 - Data structure selection (e.g., choosing XML for particular problems, relational for others); and
 - Modelling language specific guidelines (e.g., usage restrictions on XML Schemas);
- Training:
 - General data modelling principles and practices;
 - Data mapping and translation; and
 - Data model examples and use cases;
- References:
 - Lists of metadata/schema registries, e.g., DoD Data Service Environment 2.0 (DSE 2.0) [25]; and
 - Roster of data modelling Subject Matter Experts (SMEs);
- Tool recommendations and usage guidance.

NOTE: The term “data model” is used in a generalized sense and the principles and business rules apply equally to physical data models (i.e., schemas), logical data models, and conceptual data models.

Principle DADM-03: The use of data model guidance will improve interoperability, information sharing, and increase the value of data produced by, and available to, the Army by engendering common perspectives, technologies, and approaches in the data model development process.

Business Rule DADM-03a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data model guidance.

Business Rule DADM-03b: Army data model guidance should be used in the analysis, development, and implementation of data assets.

Business Rule DADM-03c: Data should be logically separated (i.e., decoupled) from applications by applying design and analysis guidance provided in the data model guidance.

Appendix G.2.4 provides some example data model development process guidelines.

Data modelling guidelines and principles are a subset of data model guidance that provide explicit advice, direction, instructions, methods or rules on how data models should be designed.

Principle DADM-04: Data models that are designed in accordance with Army data modelling guidelines and principles will increase the longevity, usefulness, and reusability of the data model, and will make information sharing (in both the near and long-term) easier and more effective.

Business Rule DADM-04a: Data stewards or an ADB designee shall develop, promote and maintain data modelling guidelines and principles as of part of the data model guidance.

Business Rule DADM-04b: Data models should be developed in accordance with Army data modelling guidelines and principles.

Business Rule DADM-04c: Data model (schema) design, specification, development, and fielding shall adhere to the Army Namespace Management Enterprise Solution (NMES) [21] [22].

Business Rule DADM-04d: Data models (schemas) documentation should include a very clear definition of the scope of the information represented by the data model and the mission use of that information.

Business Rule DADM-04e: Data models (schemas) should be designed with an anticipation of unanticipated users and future scope changes that are needed to address changing mission requirements.

Business Rule DADM-04f: Data modelling guidance for data stored in LCEs and moved into/out of DIL environment should be tailored to the limitations of the devices in the LCE and to the limitations of the DIL environment.

Examples of such tailored guidance includes:

- Scope of data model should be very explicit and unambiguous with respect to the applications and missions supported by the user/device. This will help “right-size” the data model, ensuring that the device is not storing extra, unneeded data.
- Data model shall be complete enough to support the scope and should not contain “extraneous” data.
- A logical data model may be in any data modelling format, but the physical data model (mapped/converted from logical) should be compressed/binary.

DIL/LCE

6.4 Data Quality

Data quality is a measurement or assessment of how well data meets or does not meet Army goals based on the evaluation of criteria such as relevance, accuracy, timeliness, precision, coherence, completeness, and understandability. DQM is the collection of enterprise processes and governance that ensures that enterprise data “measures up” when data quality criteria are evaluated.

Data quality directly and significantly affects the quality and effectiveness of decisions made based on that data and the actions subsequently taken based on those decisions. Therefore, the effective operation of the Army depends on high quality data.

Principle DADM-05: Effective decisions require high-quality data.

Business Rule DADM-05a: Data stewards or an ADB designee shall develop, promote, and maintain a Data Quality Management (DQM) program.

Business Rule DADM-05b: DQM processes, programs, or standards should be adopted and applied in data system design, development, and operation.

DIL/ LCE	Data quality is very important in the collection of data in LCEs and up-transmission to Command Post and Data Center CEs. LCEs should ensure the quality of data collected meets DQM requirements.
-----------------	--

Data quality is not “inspected into” data, but requires conscious attention through the chain of data entry, production, translation, transport, and use. Everyone within the Army should have awareness of or exposure to data quality importance and data quality management principles and procedures.

The best way to improve data quality is through a DQM program. A comprehensive DQM program will cover:

- Data quality training;
- Data quality metrics and measurement techniques/procedures;
- Data quality auditing procedures;
- Governance, roles, and responsibilities;
- Data quality implementation planning;
- DQM program self-correction, improvement, and optimization;
- Data quality tools, practices, and processes; and
- Data quality reporting.

A DQM program can be scoped and implemented at local or an enterprise level. Within the Army, an enterprise-wide DQM program could be complemented with tiered, localized programs.

Principle DADM-06: A comprehensive DQM program will produce and ensure high-quality data.

Business Rule DADM-06a: A DQM program should follow or adopt the DoD Guidelines for Total Data Quality Management (TDQM) as outlined in the *Army Data Framework (ADF): Data Quality Management (DQM)* [28] Section 2.1.1.

Business Rule DADM-06b: A TDQM program should establish and use a standard set of data quality dimensions to evaluate and measure data quality as outlined in the *ADF-DQM* [28] Section 2.1.2.

Business Rule DADM-06c: A TDQM program should adopt and implement the DQM best practices outlined in the ADF-DQM [28] Section 2.4.

Business Rule DADM-06d: Data quality assurance tools, mechanisms, and practices should be incorporated into system architectural specifications to ensure the quality of input and generated data and, thus, prevent low-quality data from even getting into the system. See ADF-DQM [28] Section 3.1.

Business Rule DADM-06e: A TDQM program should implement governance procedures that clearly define the roles and responsibilities for DQM as outlined in ADF-DQM [28] Section 3.2.

6.5 Data Integration

Data Integration is the process of combining data from two or more data assets and producing a single unified, consistent, and cohesive view of the combined data. Generally, the objective is to produce a set of data that represents the same information that is represented by the input data sets, though this need not always be the case. The term “data integration” also is used to refer to a data-centric strategy, approach, or architecture that (1) is designed to enable or implement an integrated, comprehensive, consistent, enterprise-spanning data deployment and management solution, and that (2) enables enterprise application interoperability. A better term for this is “Data-based Integration,” i.e., (system) integration that is based on data. Within this document, the term “data integration” will be used in the former, narrower sense.

Data integration is a concern to any data asset/system that imports data from multiple external sources. The data integration process, and the need for it, is most clearly demonstrated in Data Warehouse implementations. Data is extracted from multiple input sources, transformed and processed, and the loaded into the data warehouse, a process commonly known as ETL - Extract-Transform-Load. See *ADF: Data Warehouse* [19] for a more detailed explanation of the data integration process as part of ETL.

Data integration is more complicated than simply translating data to a common format and inserting data into a database. The data integration processor must be able to:

- Translate the incoming data to a common or canonical format (if not already in that format);
- Identify collections of data elements (i.e., “records”) in the input datasets that represent the same information (e.g., about a person);
 - If there are value differences between records representing the same information, the processor must decide whether the differences are conflicts, or whether the records really represent two (2) distinct entities²; For example, two “John Smith”s with different mailing addresses could be either a conflict of addresses or two different John Smiths; (Making this distinction is very difficult and will likely rely on probabilistic reasoning or human input.);
- Correct (automatically or through alerts to human agents) data element value conflicts between records (“cleansing” incoming data);
- Recognize and tolerate data element value synonyms (e.g., “CO” = “Colorado”);

² The term “entity,” as used in this document, refers to something of interest in the real-world, e.g., a person, location, vehicle, etc. Collections of data elements represent, stand for, or encode information about entities in information systems.

- Remove duplicate records (“de-duplicate”) so that only a single record about the same information is in the output integrated dataset;
 - Reconnect records from different datasets to the de-duplicated record, as in the following:
 - in input dataset A, a record x is connected (e.g., through a foreign key) to another record y that is about another entity (e.g., record x is about John Smith and record y is about John Smith’s automobile);
 - in input dataset B, a record x (about John Smith) is connected to a record z, about something else (e.g., John Smith’s place of employment);
 - the de-duplicated record x is placed in the output integrated dataset; the processor must then be able to connect both records y and z to x. (such that it is possible to trace relationships from John Smith’s car to his place of employment)
- This process is called “consolidation”; and
- Treat the identifier value space of the output integrated dataset as entirely independent of, and unrelated to, the identifier value space of the input datasets.

In many cases the results of a data integration process must be added to an existing dataset (e.g., a database). This may itself be treated as another data integration process: the existing dataset is just another input to the process.

Because a main objective of data integration is to recognize and remove or correct duplicative data (i.e., data representing the same information), data integration is a critical element of data quality assurance.

Principle DADM-07: A clear, robust, and well-defined data integration process is critical to ensuring data quality when data is imported into a local data asset from multiple external data assets.

Business Rule DADM-07a: Data assets/systems should adopt the data integration process outlined in the ETL process description in the *ADF: Data Warehouse* [19].

DIL/LCE	Some degree of simple data integration may take place in LCEs. Data integration can be computationally intensive so should be pushed to Command Post and Data Center CEs.
---------	---

7. Data and Services Deployment (DSD)

7.1 Enabling Information Sharing and Usage

Data and Services Deployment (DSD) is making data assets available to consumers across the Army. Enabling broad information sharing throughout the Army begins with the individual or organizational unit that owns, understands, and has the authority to “speak for” a data asset. The following items are key elements of the Army’s information sharing strategy and are presented in the order of importance:

- **Data Exposed as Services.** As data assets are identified that have information that can/should be shared with other agents, the data asset is exposed as a service to the Army/DoD (e.g., within the Global Information Grid (GIG)). Data services comply with DoD or Army data service governance documentation, e.g., CDR [19] or DSL-A [18], and apply standardized architectural patterns, such as those specified in the DSRA [14]. See Section 7.3.4 for principles and business rules associated with exposing data as services.
- **Registration to support Data and Service Discovery.** Data assets and their services are registered with the appropriate Army and DoD registries, e.g., the DoD DSE 2.0 [25]. Service interfaces are published to the registry and, thus, made discoverable to potential users of the service. Schemas and other metadata are registered with the appropriate Army/DoD registries, e.g., the DoD DSE 2.0. See Section 7.6 for principles and business rules associated with registration.
- **COI Membership.** COIs are established to address the need for a high degree of interoperability among the members and to serve as a forum for resolving specific information sharing problems. Members of a COI include the FDMs, SMEs, or PMs representing systems that need to interoperate. COIs may be established “bottom-up” by a group of system owners or data owners that need to interoperate. Or, COIs may be established “top-down” based on Joint Capability Area (JCA) [31] identification, COE Computing Environments [4], COE Control Point [4], or other enterprise-level functional planning approaches. See Section 7.5 for principles and business rules associated with Communities or Interest.
- **Information Exchange Specifications (IESs).** COIs that are organized to solve interoperability problems develop or adopt one or more community IESs that represents the data available within/among the members of the COI. The Information Exchange Specification is used as the basis for data exchange among members of the community to meet information sharing requirements. A mapping specification specifies the relationship between a schema that governs a data asset within the COI and the IES, or between the data exposed through the service interface and the IES. See Section 7.2.2 for principles and business rules associated with IESs.
- **Authoritative Data Sources.** FDMs work with Data Stewards and the Army data governance and adjudication body to identify, approve, and register Authoritative Data Sources. The data governance and adjudication body adjudicates disputes among competing claims of data authority and ensures ADSs meet quality needs of consumers. See Section 7.3 for principles and business rules associated with ADSs.

- **Unstructured Data Assets.** Huge amounts of valuable information is contained in “unstructured” data – the web pages, email traffic, chats, message boards, social network sites, imagery, video, etc., that use natural language expressions or recordings to convey information. The Army is piloting technologies and techniques for mining this data and turning it into searchable and actionable information. See Section 7.3.2 for principles and business rules associated with unstructured data.

These items summarize and provide a high-level introduction to the principles and business rules presented in Sections 7.2 through 7.6. These five (5) sections organize Data and Services Deployment principles and business rules into a collection of interrelated subject areas, as illustrated in Figure 3. Within each subject area, principles related to the topic are presented, followed by business rules derived from the principle. Collectively, the business rules provide PEOs, PMs, and FDMs system development guidance that will align their development efforts with the Army’s information sharing goals and the DoD’s information sharing objectives, and direct their data and services deployment efforts.

The five (5) sections/subject areas are:

- Data Exchange Planning and Implementation (Section 7.2);
- Data Asset Deployment Planning and Implementation (Section 7.3);
- Data Service Planning and Implementation (Section 7.4);
- Interoperability Planning and Implementation (Section 7.5); and
- Discovery and Accessibility Planning and Implementation (Section 7.6).

Data exchange planning and implementation focuses on approaches to information sharing and the development and use of IESs.

Data asset deployment planning and implementation focuses on guidance for the deployment/use of particular kinds of data assets.

Data service planning and implementation focuses on service planning, development, and deployment, particularly on the role of a data service layer within a larger SOA-based system architecture.

Interoperability planning and implementation focuses on leveraging data and service implementations within a community to facilitate and enable information sharing within and between communities.

Discovery and accessibility planning and implementation focus on guidance related to making data assets and data service visible across the Army, e.g., metadata management and registration. Registration is a process that is critical to the visibility, discoverability, deployment, and use of data and data services for information sharing.

7.2 Data Exchange Planning and Implementation

7.2.1 Anticipated and Unanticipated Information Sharing

The DoD Net-Centric Data Strategy [31] presents a vision of future DoD net-centric operations in which data is available to any and all authorized users across the DoD. The vision eschews point-to-point information sharing interfaces in favor of a “post–discover–access” model in which data providers expose their data for authorized users to discover, access, and use. This model is the same as the “World Wide Web + search engine + browser” model (plus the security/authorization needed within the DoD) in which content is posted on the web and unplanned for and unanticipated users use search engines (e.g., Google) to find information in which they are interested. This kind of unanticipated information sharing is critical to DoD operations in the future because it makes DoD information available and sharable throughout the department.

Solutions based on the post-discover-access model, however, are not well-suited to application interoperability. Application interoperability, particularly among “heavyweight” applications and in workflows, requires *anticipated* information sharing using a “compose-send-receive” model of information sharing. The compose-send-receive model involves a sender who composes a message (e.g., paper letter, email) or data package/file (e.g., database extraction) and sends it to a receiver, who receives and interprets the message. Just because data is exposed and accessible does not mean that the data can be integrated easily or clearly with other data. Data translation and data integration require deterministic, not open-ended, solutions that explicitly account for known senders and receivers of data. Therefore, Army information system development must address both anticipated and unanticipated information sharing.

Principle DSD-01: The need for Information Sharing³ may be anticipated or unanticipated.

Business Rule DSD-01a: Army data governance and architectural guidance documentation shall include strategies for addressing anticipated and unanticipated information sharing.

Within the DoD, a widely used example of anticipated information sharing that uses the “compose-send-receive” model is the Variable Message Format (VMF) [33]. VMF is a data exchange format (which has both binary and XML forms) for exchanging tactical data (e.g., observation, position, time) between combat forces and command levels.

7.2.2 Information Exchange Specifications

An IES is a document that specifies how data is to be exchanged between software applications. At a minimum, an IES specifies:

- at least one schema that governs the physical data format for the exchanged data;
- a glossary that defines the schema elements and the relationships among them; and
- the definition of extra-schema constraints governing the validity of data that conforms to the schema.

The schema and extra-schema constraints are used to validate that data that claims conformance to the IES does, in fact, conform to the IES. The glossary defines the intended meaning of the data.

³ “Information sharing” as used here in the sense of an individual act of communication, an “information sharing” event. “Information sharing” may also be used as a categorical reference to all such acts/events.

The IES may also include:

- Statements of purpose, context, scope and perspective;
- A process/activity model that establishes the information sharing requirements met (or intended to be met) by the IES, and a mapping from schema components to information sharing flows in the process/activity model;
- Semantic models for understanding (e.g., logical data model, conceptual data model, ontology, taxonomy, controlled vocabularies);
- Mapping from semantic models to the schema;
- Mapping from schema to a physical encoding format (inherent in XML Schema-XML document relationship; may need to be made explicit for compressed/binary physical formats);
- Mapping/relationship of semantic models to enterprise-level semantic models or controlled vocabularies;
- Identification of enterprise-level reusable schema components (e.g., standard data objects) used in IES schema (e.g., Intelligence Community Information Security Marking (IC ISM) [33]);
- Estimates of data exchange frequency, or expectations of the medium/method used to exchange the data; and
- Cataloging metadata, e.g., DDMS [35], to facilitate IES discovery and management.

IESs may be used in two (2) ways. The first is as a data exchange format governing a data file that is exchanged (i.e., “data-in-transit”) between two parties when information is shared using the “compose-send-receive” model. The second is as a message content format governing data made available through a data service interface when information is shared using the “post-discover-access” model. When standardized IESs are used as the basis for the message content of a data service, the need for translation and mediation is reduced.

IESs may be formally documented, published, registered, and promoted as data/information exchange standards. Many IESs have been developed, published, and are in use as standards; see Appendix E.4 *Domain-Specific-Information Data Exchange Standards* for a list of data/information exchange standards that have been developed by and are used within particular communities.

IESs serve the information sharing needs for a particular community of interoperating agents. The membership of an interoperability community may be loosely defined in the case of informal communities, or well-defined in the case of formal communities such as COIs. The membership may even be undefined in the case of general purpose IESs. Section 7.5 specifies the principles and business rules associated with interoperability within a community.

Principle DSD-02: Data models that are formalized and adopted as IESs will facilitate and enable effective information sharing within a community.

Business Rule DSD-02a: IESs shall be formally documented in accordance with Army policies, templates, and other requirements governing IESs. At a minimum, the IES shall include a schema, the definitions of schema elements and the relationships among them, and the definition of any extra-schema constraints governing the validity of data that conforms to the schema.

Business Rule DSD-02b: The data models upon which IESs are based shall follow, adhere to, or comply with Army Data Model Guidance (see Section 6.3).

DIL/LCE	Business Rule DSD-02c: The data models (schemas) in IESs intended for use in DIL environments shall include a mapping/conversion to a compressed, binary physical exchange format unless a network impact study is conducted and a waiver is obtained.
----------------	--

Reuse of IESs is key to reducing the proliferation of data formats.

Principle DSD-03: Reusing published IESs facilitates interoperability across the Army.

Business Rule DSD-03a: Communities should pursue the adoption of an IES in the following preferential order:

- Adopt and use a published IES or Information Exchange Standard (see Appendix E.4) as-is;
- Research metadata/schema repositories such as the DoD DSE 2.0 for a data model(s) relevant to the community's information sharing requirements and adopt that data model(s);
- Modify and adopt a published IES or data model discovered in a metadata repository; and
- Develop and adopt a community-specific IES.

The reuse of IES in the design of data services as the description of the service message content format will reduce the need for mediation.

IESs define the data exchange formats for interactions in the End-State Information Sharing Framework illustrated in Figure 7.

The *Rules for Cross-Cutting Capability (CCC) Information Exchange Specifications (IES) in Interface Specifications* [23] provides additional, more specific, business rules associated with adoption and use of IESs.

The DSRA Information Architecture [18] defines a pattern called *Creating Data Exchange Specifications* that may be applied to the development of an IES. AR 25-1 [12], Section 5-2.e, specifies Army policy regarding IESs.

NOTE: Appendix G.2.2 describes the processes associated with IESs.

7.3 Data Asset Deployment Planning and Implementation

A data asset is data that has been enabled with some mechanism to make data within the asset available to consumers in the Army. The mechanism may be as simple as a procedure for exporting data to a file and exchanging the file, or as sophisticated as a web service integrated as part of a SOA deployment. The data asset deployment "base case" is making a database available to Army consumers, e.g., by providing and publishing a data service that provides access to the database. There are many different kinds of data assets throughout the Army, however. Planning for the deployment and use by consumers of the data assets will depend, in part, on the nature, preparation, and purpose of the data asset. This section presents the guidance for the following kinds of data assets found on Army networks:

- Authoritative Data Source (ADS)
- Unstructured Data Asset
- Cloud-based Data Asset
- Enterprise Resource Planning (ERP) System-based Data Assets

DIL/LCE	<p>LCEs shall not host data assets of these types. Services or applications in DIL environments may access data from these kinds of data assets.</p> <p>Devices in LCEs may be considered “data assets” insofar as they are data collection devices that obtain and upload data that is of value (sometimes critical value) to the Army (e.g., sensors). See “Data Production/Acquisition Sources” in Figure 8. They are not, however, “data assets” in the sense of being a persistent data asset that can be discovered and accessed by consumer across the Army.</p>
----------------	---

7.3.1 Authoritative Data Sources (ADS)

An approach adopted by the Army to address data quality problems (e.g., conflicting or outdated data) was the definition and fielding of ADSs. An ADS is an Army data asset recognized by a governing data authority as the single authoritative source for a particular kind of information (i.e., a “data need”). ADSs will reduce or eliminate the problem of conflicting data by designating the data asset as the official Army source for that kind of data; they also provide an explicit focus point for data quality efforts.

Principle DSD-04: Timely, effective, and accurate decision-making depends on timely and accurate information; timely and accurate information depends on the availability and quality of Authoritative Data.

Business Rule DSD-04a: If a data asset contains information that may support: (1) an Business Enterprise Architecture (BEA) end-to-end process [37] or (2) a JCA capability [31], then the data asset should be submitted to and registered with the DoD DSE 2.0 for consideration and certification as an Authoritative Data Source. The submission of a data asset for consideration may entail adjudication of competing claims of authority or jurisdiction.

Business Rule DSD-04b: A data asset that has been certified as an ADS by the appropriate designated body shall be maintained in accordance with policies and procedures that govern ADSs.

Business Rule DSD-04c: When timeliness/currency of data is important, real-time access to ADSs is preferred over non-ADS sources.

Note that ADSs are both 'registered' and 'approved' within the DSE 2.0. A 'registered' ADS is an ADS that has been entered into the DSE 2.0, regardless of its approval level. An 'approved' ADS is an ADS that has been entered into the DSE 2.0 and is in the 'approved' state and therefore considered to be an officially recognized ADS.

Information consumers should not have multiple choices from which to obtain certain kinds of information, e.g., soldier medical records or spare part inventories. The data that represents this information should be published in one place (e.g., in an ADS) and made available across the Army.

Principle DSD-05: Information (of a given type) that is available from a single source (rather than multiple sources) will reduce the possibility of conflicting information and increase the trustworthiness of the information.

Business Rule DSD-05a: An ADB designee should analyze, evaluate, and plan the content of ADSs across the Army such that specific kinds of information (i.e., “data needs”) are not provided by multiple sources.

DIL/LCE

LCEs may host and maintain cached ADS data as long as the cache is refreshed on a regular basis when connected.

An ADS is shown as a kind of data asset in the End-State Information Sharing Framework previously illustrated in Figure 7.

AR 25-1 [12], Section 5-2, specifies Army policy regarding ADSs.

Appendix G.2.1 describes the processes associated with ADSs.

7.3.2 Unstructured Data Assets

The term “unstructured data” is typically used to describe data content that is expressed as the written form of natural language rather than as fields or data elements associated with a data model. It is also used to refer to digital image, video, or audio recordings. Almost all World Wide Web content, including social networking sites, is, or contains, “unstructured data.” The meaning of unstructured data typically requires human interpretation and reasoning, though new technologies such as IBM’s Watson [37] have demonstrated the ability to “understand” natural language text.

Unstructured data assets are valuable sources of information. Statistically-based Artificial Intelligence (AI) techniques may be used for analyzing unstructured data, identifying “entities” in the data (e.g., a person, a physical location), and then tagging that data with metadata that identifies the entity. The metadata can then be connected to structured data about the same entity, thus increasing the semantic richness of the unstructured data. These techniques transform unstructured data into semi-structured data. Semi-structured data is grouped with unstructured data in the AIA for the purpose of explanation even though its “structured-ness” lends itself to some automated processing and sharing by software applications. Figure 8 illustrates a spectrum of structured-ness of different kinds of data assets.

Principle DSD-06: Unstructured data is a valuable source of information.

Business Rule DSD-06a: Data stewards or an ADB designee should (1) recommend technologies for the semantic analysis of unstructured data (e.g., written prose and recordings) and (2) develop an Army-wide strategy and guidance for harvesting and leveraging the information in unstructured data.

Business Rule DSD-06b: A catalog of unstructured data assets (or locations) that contain information of value to the Army should be created and maintained. The DoD Discovery Metadata Specification (DDMS) [35] should be used to annotate/tag each unstructured data asset in the catalog to facilitate discovery.

Business Rule DSD-06c: Unstructured data assets should be analyzed and tagged in accordance with Army strategic guidance.

Both structured and unstructured data may be maintained and managed in a “data cloud” (see *ADF: Data Aspects of Cloud Computing* [37], Section 3.1) where the data undergoes general indexing, categorization, and entity extraction to enable user search for information over a broad collection of data. Rather than integrating multiple distinct data assets into a single canonical form, hybrid or multi-structure “data cloud” data assets use indexing, categorization, and entity extraction that spans and connects multiple containers of differently-structured data.

Unstructured data is shown as an example data asset in the End-State Information Sharing Framework illustrated previously in Figure 7.

NOTE: Appendix G.2.3 describes the processes associated with unstructured data and provides references to additional information and resources concerning unstructured data.

7.3.3 Cloud-based Data Assets

Deploying data assets to a Cloud Computing Environment (CCE) offers both data providers and data consumers valuable capabilities and features (when compared to conventional computing environments), such as (from [40]):

- *“On-Demand Self-Service.* A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.
- *Broad Network Access.* Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
- *Resource Pooling.* The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.
- *Rapid Elasticity.* Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
- *Measured Service.* Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.”

Consistent implementation and use of cloud computing technology will benefit the Army and all users of the technology.

Principle DSD-07: The use of cloud computing implementation guidance will improve interoperability, information sharing, and increase the value of data produced by, and available to, the Army by improving the reliability and availability of data.

Business Rule DSD-07a: Data stewards or an ADB designee shall develop, publish, promote, and maintain cloud computing implementation guidance in accordance with Federal CIO and National Institute of Standards and Technology (NIST) guidance.

Business Rule DSD-07b: Cloud computing implementation guidance shall be used in the design and development of cloud-based systems; in particular, the following guidance documentation shall be used:

- Army Data Framework - Data Aspects of Cloud Computing [37].

Business Rule DSD-07c: A migration/deployment plan shall be developed to guide the movement/deployment of data or data services to a CCE. The plan should follow, or be compatible with, the guidance provided in the *Federal Cloud Computing Strategy* [42] and Appendix G.4. (See also Section 2.4.1 of *Army Data Framework - Data Aspects of Cloud Computing* [37].)

Many design decisions involved in deploying data to a CCE are the same as those in any other environment. For example, the storage implementation structure for the data must still be suitable for the applications that will use the data:

- Relational for transaction management;
- Dimensional for data warehouses and business intelligence; and
- Key-value storage for “Big Data”⁴.

In other words, just because Hadoop is popularly associated with cloud computing technology does not mean that it is always an appropriate paradigm for cloud data.

Similarly, data services for accessing cloud data must adhere to the same interface standards and security measures as non-cloud data services, such as well-defined application programming interfaces (APIs) and role-based access.

This does not imply that real-world constraints and performance characteristics of cloud computing technology do not impact the design decisions involved in implementing cloud-based data assets. The functional requirements of the application must be integrally balanced with the limitations and characteristics of cloud technology, e.g., throughput of the physical machine that is hosting multiple Virtual Machines (VMs).

Principle DSD-08: The design decisions on the data storage implementation paradigm (e.g., relational, key-value, dimensional) and access methods (e.g., data services) for cloud-based data assets depend more on application/usage requirements than on Cloud Computing Technology.

Business Rules DSD-08a: Cloud data storage and access should be designed to meet, and be suitable for, application usage requirements and leverage cloud computing benefits while also addressing cloud computing technology performance constraints and limitations.

Leveraging the capabilities of cloud technology makes the notion of *Data as a Service* (DaaS) possible.

Data deployed to the cloud is hosted on infrastructure owned by the cloud service provider, which is typically shared among multiple tenants and spread across multiple locations. This shared hosting model in the cloud introduces risks that need to be considered and mitigated, including:

- Security:
 - How is the data (including backups or archives) protected from other cloud tenants and external intruders, particularly in newly-provisioned or de-provisioned resources?
 - What access controls are used to protect data?
 - Where is the data physically located and how is the physical facility protected?

⁴ “Big Data” can be described as “...data sets that [are] so large and complex that they become awkward to work with using [conventional] database management tools... current limits are on the order of petabytes, exabytes and zettabytes of data” [62]

- What happens to residual/remanence data when data is migrated from or moved off of the cloud (i.e., decommissioned)?
- Legal:
 - Who owns the data?
 - Who is liable for data loss or compromise?

Security and legal considerations are, therefore, more critical and important to address for data in a CCE and command a larger percentage of design attention than data deployed in conventional environments.

Principle DSD-09: Data security and legal concerns are of greater significance when data is deployed in/to a Cloud Computing Environment (CCE) when compared to conventional computing environments.

Business Rule DSD-09a: A security plan shall be developed in conjunction with the movement/deployment of data or data services to a CCE. The security plan should address the security and privacy challenges presented in *Challenging Security Requirements for U.S. Government Cloud Computing Adoption* [43] and the security/legal considerations presented in Appendix G.4.

Business Rule DSD-09b: Data shall be deployed to a CCE In Accordance With (IAW) the security requirements and guidance provided by the Federal Risk and Authorization Management Program (FedRAMP) [42].

Under FedRAMP, government agencies can leverage pre-authorized Certification and Accreditation (C&A) packages and pre-approved applications, which will reduce the duplication of effort to certify the same application numerous times across the government.

Principle DSD-10: A clear legal contract between a cloud service provider and cloud service consumer protects both the provider and consumer.

Business Rule DSD-10a: A clear, unambiguous Service Level Agreement (SLA) or legal contract shall be prepared and signed by cloud service consumer and cloud service provider.

Additional guidance and resources on cloud security can be found in the *Army Data Framework - Data Aspects of Cloud Computing* [37].

Additional guidance on implementing or migrating data (and data assets) to a CCE is provided in Appendix G.4.

7.3.4 ERP-based Data Assets

Enterprise Resource Planning (ERP) is a business management approach that encompasses a broad set of practices supporting company operations (e.g., finance, purchasing, logistics, human resources, and payroll) from a holistic, integrated, and enterprise-wide perspective. ERP systems are large software systems that integrate and manage company operations and functions; they provide seamless integration of end-to-end processes across functional areas, offer improved workflow, and institutionalize standard business practices. A typical goal of ERP initiatives is to replace aging, stove-piped information systems with a single enterprise-wide solution that provides extensive functionality and “point-and-click” access to real-time operational data across the business.

Although ERP systems are intended to holistically manage the operations of an entire enterprise, an enterprise as large as the Army can (and does) host several ERP systems. When it comes to interoperability of Army systems, an ERP system is treated like any other single application when it comes to data exchange, information sharing, and interoperability.

Data management (particularly Master Data Management (MDM), see Section 7.5.1) is key to the operation of ERP systems and integration of enterprise processes. As such, ERP systems are a rich and deep source of data that is valuable to the Army. So making that data available to Army consumers outside the ERP system environment is an important aspect of ERP implementation and deployment.

Principle DSD-11: ERP systems are the same as other software systems/applications in the Army when it comes to data exchange and information sharing: ERP systems interoperate with other Army systems and may serve as a data asset for consumers across the Army.

Business Rule DSD-11a: ERP systems shall make their data available to consumers as appropriate to support interoperability with other systems, particularly Master or Authoritative data,

Business Rule DSD-11b: ERP systems should be implemented IAW the guidance provided by the *ADF: Enterprise Resource Planning* [45].

7.4 Data Service Planning and Implementation

7.4.1 Data Services

A “service” within an SOA-based system is a callable interface to an application that provides a function or capability to the calling agent (e.g., a consumer or client). A data service is a service that provides access to a data asset or a data management function/capability. Within the Army’s SOA-based infrastructure development, a *data services layer* is the part of a SOA framework that “sits over” the data assets on the network and provides access to those data assets; see the middle Services Layer illustrated in Figure 6 and the middle layer of the End-State Information Sharing Framework previously shown in Figure 7.

Once a data service is deployed on a network, the service interface specification is published to a public registry (see Section 7.6). There, consumers can discover the service, download the service specifications, and develop client applications to call and make use of the service. Data services are a key ingredient for meeting the DoD information sharing objectives because they make data *accessible* to unanticipated users.

Principle DSD-12: Data services contribute to meeting unanticipated information sharing requirements. Exposing data via data services makes data available to and accessible by unanticipated, authorized users.

Business Rule DSD-12a: Data of value to the Army shall be made available to authorized consumers in the Army via data services.

Data services are depicted as “plug” icons in the End-State Information Sharing Framework previously illustrated in Figure 7.

While data services provide access to data assets, a data service itself is considered a data asset, as illustrated in Figure 8.

Authorized consumers are data service consumers who have been authenticated according to applicable security requirements and have sufficient permissions to use the service. See Section 8 for more information on secured availability.

Data Services fall into three (3) broad categories: Enterprise Data Services, Provider Data Services, and Data Management Services. (See Figure 7, Consumer Data Services consists of both Enterprise Data Services and Provider Data Services.)

- Enterprise Data Services are global-use services that provide a unique service to consumers across the Army. An Enterprise Data Service is developed and maintained by a single organization.
- Provider Data Services are distributed throughout the Army; a Provider Data Service may be an instance of a standard service, or it may be a unique offering of the provider. For example, a “data access” service may use a standardized service interface (e.g., DSL-A Retrieve Service), but there may be many instances of data access services throughout the Army based on the same service standard.
- Data Management (or “Back Office”) Services are infrastructure services that are not generally accessible to consumers but are necessary to provide complete data management capabilities.

Service reuse is a key value proposition of a SOA-based system design approach.

Principle DSD-13: It is better (e.g., more cost effective) to reuse existing services than to develop a new service.

Business Rule DSD-13a: Before the development of a data service is undertaken, service registries should be searched for both existing, fielded services, and services that are under development that could meet the requirements of the required data service. The DSE 2.0 is the primary service registry that should be researched.

Business Rule DSD-13b: Where a service exists that meets the requirements of the required data service, the existing service shall be adopted and used and a new data service shall not be developed. If multiple services exist that fulfill a capability need, a data service shall be chosen and adopted in the priority order presented in Table 3.

Business Rule DSD-13c: Where a data service (or services) exist that partially meet the requirements of a required data service, the owners of the required data service should engage the owners/maintainers of the existing data service to request a change in order to meet the requirements of the required data service. If a good-faith effort to change the existing data service fails in a reasonable length of time, then the required data service should be developed.

A Service Interface Specifications (SIS) is a software specification document that specifies how a service is to be constructed and what it is supposed to do. Data service standards are service interface specifications that have been certified, approved, or authorized by a governance body for use in the development of service implementations. They are published to a publicly available resource (e.g., a registry/repository) where they may be discovered and downloaded by potential users. A standardized service interface specification includes an explanatory document, a formal service interface definition (e.g., a Web Services Description Language (WSDL), and, optionally, a message format schema (e.g., an XSD). A data service may be either SOAP-based [45] or RESTful [47]; the messaging approach adopted/used in the service will be specified in the service interface specifications. The value in reusing standardized service specifications is that much of the service design work has been done and the service specification already adheres to required data service governance, such as service structure,

element naming, and security handling. In addition, implementation tools are often available to support the implementation of the specifications.

Principle DSD-14: If no service exists that fulfills a capability need, it is better (e.g., more cost effective) to use existing service interface specification standards for implementing the service than to implement a service with a unique, localized interface.

Business Rule DSD-14a: If no web service exists that meets the requirements of the proposed data service, service and metadata repositories shall be searched for published/standardized data SIS that can fulfill the capability need. If no suitable SIS is found, a new SIS shall be developed and submitted to appropriate service and metadata repositories. If a suitable SIS is found, the specifications should be adopted and implemented as published; if the SIS only partially meets the requirements of the proposed data service, the authors of the SIS shall be engaged to request a change to the SIS; if a good-faith effort to change the SIS fails in a reasonable length of time, then a new SIS should be developed (or the partially suitable SIS extended) and submitted to appropriate service and metadata repositories.

Examples of data service standards include the Intelligence Community/Department of Defense (IC/DoD) CDR specifications [19] and the DSL-A specifications [18].

The flow and dependencies that are part of these principles and business rules are illustrated in the flow chart presented in Figure 9. A description of a data service development process is presented in Appendix G.2.5.

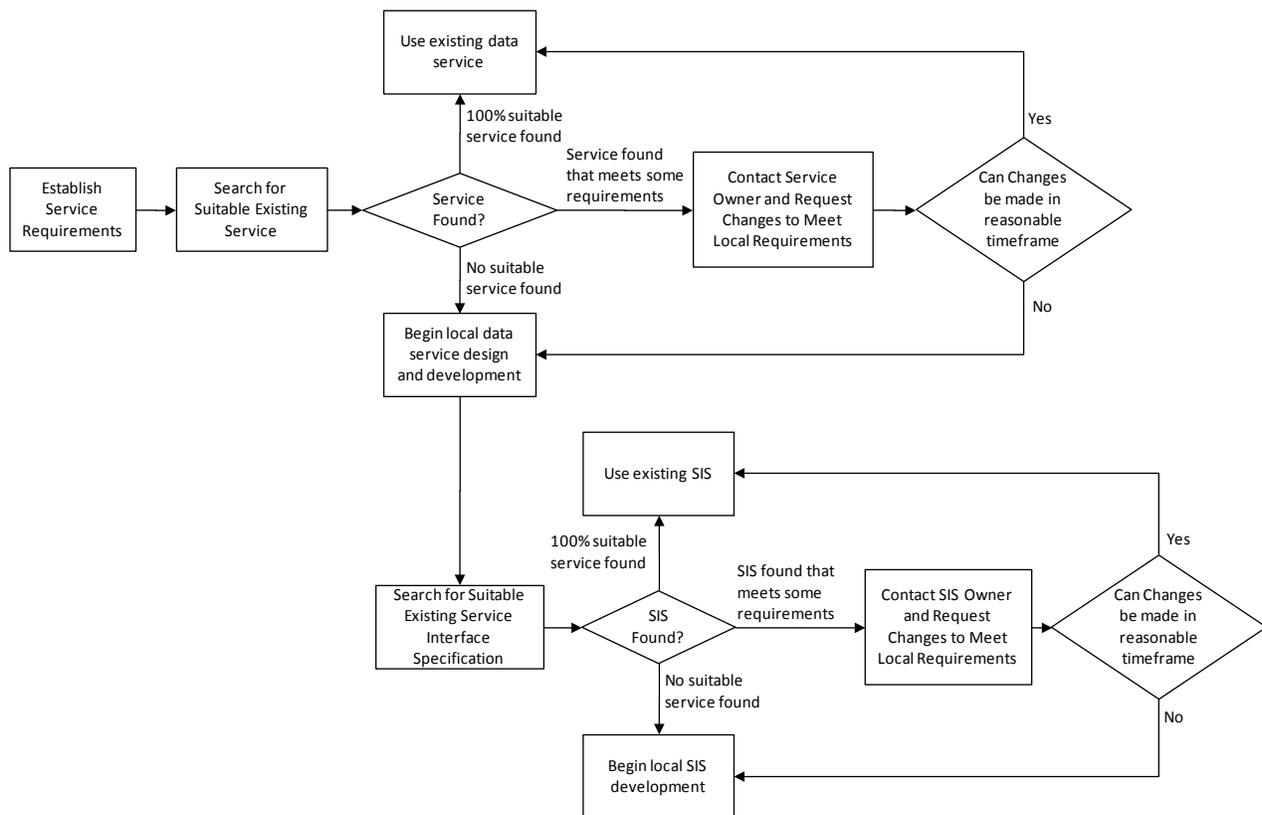


Figure 9: Service Reuse and Development Flow Chart

Appendix G.3 describes the processes associated with service development and fielding. In parallel with developing and deploying provider data services, Enterprise Data Services will be developed by an enterprise infrastructure team, such as the ones developed at the DoD level, including:

- GIG Content Delivery Services (GCDS) optimize the delivery of Web content and applications by distributing them on globally deployed servers. End users access content via the geographically closest enclave, resulting in increased connection speed and minimal download times.
- Enterprise File Delivery (EFD) is a standalone lightweight application used to synchronize large file storage between geographically separated sites, helping to achieve Wide Area Network bandwidth savings.

The development of new Enterprise Data Services should follow the same general process as for a Provider Data Service in that reuse is a goal.

The Army should search across the DoD and Army to identify these services, find the best of breed and adopt and maintain a single implementation.

Principle DSD-15: Some data services offer a capability that is best developed and provided by a single service that is used by consumers across the Army.

Business Rule DSD-15a: Data stewards or an ADB designee shall be responsible for Army service portfolio management. This body (or bodies) shall research, identify, design, develop, deploy, and manage Enterprise Data Services that provide single-source capability to consumers across the Army.

Business Rule DSD-15b: Data services should be considered for suitability as an Enterprise Data Service. If suitable, the data service should be brought to the attention of the Army's service portfolio management body.

DIL/LCE	Applications/devices in LCEs in DIL environments may use any Army data service as appropriate given the constraints of the LCE and DIL environment. However, services hosted at CEs "closer" to the application/device in DIL environments are likely to be of greater value to the application/device than a similar enterprise-level service because of latency/response time.
----------------	--

7.4.2 Data Service Guidance

In the development of a data service, there are many design choices that could result in non-uniform or hard-to-use service interfaces. Data services development is a repeatable process – many of the development steps are the same regardless of the type of service (e.g., incorporating security controls). Data service guidance can improve the development process by standardizing the steps and tools necessary to create a service. The consistent use of data service guidance in the development of data services across the Army will yield more consistent, robust, and reusable data services.

Data Service Guidance is the collection of procedures, methods, best practices, recommendations, and subject matter expertise that support data service design, development, and implementation. Data service guidance includes:

- Standardized development process, e.g., the DSL-A Developers Guide [37], Appendix G.2.5;
- List of data service standards and supporting standards, like security standards;

- Use of namespaces;
- Standardized error handling; and
- Support tools, such as those that support data services development, testing and deployment (e.g., Common Data Service Framework (CDSF) [49]).

Principle DSD-16: The use of data service guidance will ease the process and reduce the cost of both (1) creating and deploying data services, and (2) using data services.

Business Rule DSD-16a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data service guidance. The same body, or an allied body, should develop, test, and promote service development support tools (e.g., CDSF) and resources.

Business Rule DSD-16b: Army data service guidance should be applied in the design and development of data services.

Business Rule DSD-16c: The design and development of a data service should:

- follow the Army's data service development process;
- adhere to the Army NMES [21] [22]; and
- incorporate instruments or monitors to track performance and usage of the data service.

Business Rule DSD-16d: In DoDAF architectures, data services should be documented as a SvcV-4 Services Functionality Description view.

DIL/LCE

Business Rule DSD-16e: Data service guidance shall take into account constraints of LCEs and DIL environments for the design of services to be deployed on LCE devices or in DIL environments.

Principle DSD-17: Data services that are designed and implemented in accordance with Army data service guidelines to be as “future proof” as possible will increase the longevity and reusability of the service.

Business Rule DSD-17a: The design of a new data service should anticipate other potential users of the service (e.g., outside the known consumers of the service) and consider potential future uses. Data services should not be (effectively) a point-to-point service for a particular requirement. The data service should be generic and/or flexible so that it can be reused. This is what is meant by “loose coupling.”

Business Rule DSD-17b: When a data service is changed, backward compatibility should be maintained so that existing clients of the service will not be affected.

Business Rule DSD-17c: Data services should be designed to be scalable to handle more users than the number currently anticipated.

DIL/LCE

Business Rule DSD-17d: The design of services to be deployed on LCE devices or in DIL environments may be exempted from general data service guidelines with a justification based on LCE or DIL environment constraints. Such services shall comply with any guidance specific to the LCE or DIL environment.

Principle DSD-18: Monitoring of the operation and performance of a data service will ensure that the service meets user expectations and serve as feedback to improve the data service design and development process.

Business Rule DSD-18a: The performance of the data service should be monitored to ensure performance remains within acceptable limits to all users.

Business Rule DSD-18b: PoRs should monitor data consumption actions and feed statistics back to PMs, CIO/G-6, and the ADB to improve AIA and development processes.

DIL/LCE	Business Rule DSD-18c: The design of services to be deployed on LCE devices or in DIL environments may exclude monitoring capabilities with a justification based on LCE or DIL environment constraints.
----------------	--

Enterprise Service Management (ESM) products and solutions often include service monitoring functionality.

NOTE: Appendix G.3 describes processes associated with and that support consistent data service development.

7.5 Interoperability Planning and Implementation

Interoperability planning and implementation requires a systems engineering perspective over a set of inter-related and collaborative systems, which establishes the scope of interoperability solutions. IESs naturally play a role in interoperability solutions and, therefore, play a role in some of the topics addressed in the following subsections.

7.5.1 Master Data Management (MDM)

Master data represents information that plays an essential or key role in the operation of a business. This information is typically non-transactional information about business entities such as customers, products, inventory, or suppliers, where the correctness and currency of the information is critical to business operations. Master data is typically shared and used by different software applications across the enterprise, often as part of transaction processing. Master data provides a continuity and consistency of knowledge throughout the enterprise.

MDM is the set of processes and tools that ensure that master data is effectively controlled, updated, and used within and throughout the software systems used by the enterprise. MDM “sits on top” of and leverages many existing technologies:

- Data warehouse technology to maintain a master data “hub” containing the official, definitive, or authoritative version of the master data;
- Data quality principles and metrics to ensure that master data is correct and up-to-date (see Section 6.4);
- Data integration to support duplicate entity recognition and de-duplication during the data warehouse ETL processes (see Section 6.5); and
- Data mapping and translation to convert data from source formats to the canonical master data format (see Section 7.5.3).

In a sense, MDM is a practice that is woven into software systems to provide a global data quality assurance capability for data deemed most critical to the enterprise. It involves:

- the explicit recognition of sources of master data, i.e., the systems that create it;
- explicit identification of systems that use (and possibly modify) master data; and

- the insertion of data quality assurance and integration processes into general data management capabilities built into enterprise software systems that deal with master data.

The key is that master data is always correct and current, which is the ultimate goal of MDM.

Authoritative Data, as defined by the Army (see Section 7.3.1), is a subset of Master Data; it is data that has gone through the approval process and, once approved, designated as the correct or authoritative (trusted, reliable, and accurate) version of the data. Master data does not inherently require the data to be considered authoritative.

The designation of data as “master” data and maintaining the quality of the data enables that data to serve as an integrating mechanism across enterprise software systems because it is always correct, current, and commonly understood. Master data enables interoperability among systems.

Principle DSD-19: Master data is an important mechanism for integrating enterprise software systems. Master data management ensures that master data is always correct and current.

Business Rule DSD-19a: Master data management processes should be included in interoperability architectures as described in *ADF: Master Data Management* [50].

Business Rule DSD-19b: Master data management processes shall incorporate Unique Identifiers (UID) for significant Army assets as directed by DoD Directive 8320.03 *Unique Identification (UID) Standards for a Net-Centric Department of Defense* [51].

The use of UIDs for identifying discrete, significant assets on an Army-wide basis will facilitate master data management by providing a “linchpin” identifying value for reconciling and integrating data about an asset.

DIL/LCE	Business Rule DSD-19c: Applications/services in LCEs may use master data but should not be responsible for any master data management functions.
----------------	--

7.5.2 Community-based Information Sharing

There is no single “blanket” solution to information sharing that can be introduced and used across the Army. Information sharing solutions must start within well-understood, managed communities. Once an information sharing solution in a community is operating effectively, then it can be expanded to encompass other, wider communities.

Interoperability communities can be an informal, loosely organized group of members or a formal group that is organized as a COI, where a member is a system, service, or data asset that is coupled with a human representative. Interoperability communities can be described by the following properties:

- Members of the community/COI share information frequently in collaborative pursuit of a mission;
- There is a describable body of information that is the primary subject, interest, or domain of responsibility of the community/COI; and
- The data that represents this body of information among the members of the community/COI, and the exchange of data to entities outside the community/COI, is the collective responsibility of the community.

Community-based interoperability solutions lead to the following principles and business rules.

Principle DSD-20: COIs are the basis for anticipated information sharing in the Army, and for defining and meeting interoperability requirements.

Business Rule DSD-20a: Information producers and consumers that regularly share information should join (or form) and participate in COIs.

Membership in a COI need not be restricted to Army stakeholders, but may include other DoD or Joint stakeholders as well. Appendix G.1 describes the processes associated with COI membership, formation, and operation.

Meeting interoperability requirements and developing interoperability solutions are more effectively developed on a small, local scale rather than on enterprise-wide scale.

Principle DSD-21: Interoperability/collaboration is most effectively achieved, and an interoperability solution is most effectively designed and implemented, within a small community. Interoperability can be effectively scaled from small communities to larger communities.

Business Rule DSD-21a: COIs should focus on data exchange among systems within a COI before considering data exchange within a broader community.

Business Rule DSD-21b: COIs should have an Information SME who is responsible for knowing the data assets within the COI, the information they contain, and the relationships among them. This role is similar to that of an FDM with a scope of responsibility that covers the COI.

Scalability of the interoperability solutions is achieved through the “community of communities” concept, in which a member community is represented in the higher community by the IES used by the member community.

A COI is shown as dotted box on the left side of the End-State Information Sharing Framework illustrated in Figure 7.

Interoperability requirements – the requirements that must be met by an IES for effective use of the IES for information sharing – must be understood and documented. Interoperability requirements are obtained by analyzing the potential interactions (i.e., the information sharing events) between members of the community. The result is the documentation of the data exchange pathways (for information sharing events) among members of the community.

Principle DSD-22: Unambiguous interoperability between highly interactive applications (particularly those requiring data translation and data integration) requires overt, formal, and precise specification of data exchange pathways and information content of the exchanged data (i.e., the description of the anticipated information that is being shared).

Business Rule DSD-22a: COIs shall create and maintain a DoDAF AV-2 Integrated Dictionary and should create and maintain a DIV-2 Logical Data Model.

The AV-2 documents the “common vocabulary” of COI and the DIV-2 documents the abstract, logical view of the data exchanged among members of the COI.

Business Rule DSD-22b: Applications/systems that interoperate frequently with other applications, particularly within a COI, should have an explicit documentation of the interoperations (e.g., resource flows or “data exchange pathways”) between the applications/systems. COIs should develop and document these interoperations using sets of DoDAF models [19]; the sets of models that document the interoperations are shown in Table 4; either or both sets (e.g., Business Process View or System Interaction View) should be developed.

Table 4: Levels of Interoperation Description

Level of Detail	DoDAF Models	Description
Business Process View	OV-2: Operational Resource Flow Description	Information Resource Flows along Need-lines between operational activities of members of COI.
	OV-3: Operational Resource Flow Matrix	Detailed description of Information Resources that flow.
	OV-5b: Operational Activity Model	Process/activity model of interrelated activities of members of COI.
System Interaction View	SV-1: Systems Interface Description	Describe the physical solution interface between systems (including ports, formats.)
	SV-2: Systems Resource Flow Description	Precise specification of a connection between systems
	SV-4: Systems Functionality Description	Physical equivalent of OV-5b.
	SV-5a: Operational Activity to Systems Function Traceability Matrix	Maps system functions to business processes.
	SV-6: Systems Resource Flow Matrix	Physical equivalent of OV-3.

Information is shared between collaborating agents via the exchange of data between systems used by the agents. IESs specify how data is to be exchanged to meeting information sharing requirements.

Business Rule DSD-22c: COIs should adopt, or develop and publish, one or more Information Exchange Specifications (IES) that represents the information available within the COI. COIs should adopt and use published IES standards where/when possible. The COI Information SME is responsible for the development of and is the custodian of the IES.

Business Rule DSD-22d: The physical structure and information content of exchanged data shall be documented and governed by an IES.

DIL/LCE	All interoperability/COI business rules apply to LCEs.
----------------	--

7.5.3 Translation and Mediation

Translation and mediation are critically dependent on data modelling techniques and tools. Physical data models (schemas) specify what data “looks like” in a data asset and in exchanged files/messages, and physical data models are the basis for specifying the mapping and translations between data assets. *Information sharing* is a communication act that has four (4) components:

- A subset of data is extracted from a data asset corresponding to the information to be shared/transmitted;
- The data is transmitted to a receiving application/data asset as a file or message;
- The received data is transformed from the received format to the local data model format of the receiving application/data asset, where the transformation maintains, to the highest degree possible, the information content (i.e., semantics) of the received data; and
- The transformed data is imported into and integrated with the target data asset.

The process of transforming data from one format to another while maintaining the information content is called *translation*, just as transforming an English language sentence into French is called “translation.”

The relationships between data models (e.g., a local physical data model and an IES) are defined by *mapping specifications*. Mapping specifications govern and drive the data transformation action; COTS tools are available for developing, testing, and documenting mapping specifications, and automatically generating translation code based on the mapping specification.

Principle DSD-23: The formal specification of the relationship between two (2) different data models (i.e., schemas) or IESs is necessary to understand, monitor, and maintain consistent information content (i.e., semantics) of data during a data transformation process.

Business Rule DSD-23a: Mapping specifications should be developed and published that define the relationship between local schemas and the IESs used to share information with collaborating partners. A “local schema” may be a schema published with a data service interface (i.e., “export schema” or “service schema”) or the schema of a data asset. The mapping specification shall be detailed enough to unambiguously define the data transformation process, and shall identify semantic gaps that result from the mapping.

Business Rule DSD-23b: Formal mapping specifications shall be used to govern the transformation of data from a format conforming to one schema to format conforming to another schema.

Mapping and data transformation apply to structured data.

While mapping provides a significant value in explicitly and overtly representing the relationship between two (2) physical data models, it also provides – as perhaps a more significant value – a deterministic mechanism for tracing semantic errors in data exchanges. If “strange” data (e.g., out of range, unexpected, or wrong) is received by a consumer, the mapping specifications are a mechanism for tracing the translations that took place between the source and the receiver and determine the cause (e.g., misuse of the IES by the sender) of the semantic error.

Mediation is the process in which data is translated from its original schema format to a schema format more suitable for the receiver through a mediating agent. Mediation involves a “third party” neutral mediating format (e.g., one governed by an IES) that acts as an intermediary

between the sender and receiver; two (2) translations are involved in exchange of data when a mediating form is used. Mediation may involve a sequence of transformation/translations stages. A mediation service is illustrated in Figure 7, the End-State Information Sharing Framework; it shows data being drawn from three (3) difference data assets, translated and integrated into a mediating format (i.e., as defined by an IES) by the service, and delivered to consumers in an IES format.

Sequential transformation/translations may introduce syntactic or semantic gaps or misinterpretations that could negatively impact the fidelity of the information being shared. Careful analysis of mapping gaps is required to preserve information fidelity across multiple translations.

NOTE: The DSRA Information Architecture [18] defines several patterns that described the mediation and translation process: *Data Set Transformation*, *Data Mediation*, and *Data Abstraction*.

NOTE: Appendix G.2.5 describes the processes associated with mapping, translation, and mediation of data.

7.6 Discovery and Accessibility Planning and Implementation

Discovery and Accessibility planning and implementation directly address the information sharing objectives of Visibility, Accessibility, and Understandability. Metadata management contributes to the accessibility and understandability of data. Registration makes a data asset visible (i.e., discoverable).

7.6.1 Metadata Management

Metadata is data that is “about” some object and describes features or characteristics of the object with respect to some purpose. For example:

- Card catalogs in libraries contain metadata describing books for the purpose of finding particular books.
- Databases in motor vehicle departments contain metadata about automobiles for the purpose of tracking and licensing the vehicles.
- Schemas are metadata about databases for the purpose validating and understanding data in the database.

Precisely defining “metadata” is difficult because the meaning of “metadata” is relative and recursive: it is applied to something that is really “about” something else, and metadata may itself have metadata, which may have metadata. Metadata can be broadly described as falling within three main categories, as illustrated in Figure 10. These categories are a hybrid of several extant definitions; see *ADF: Metadata Management* [51] for a description of these definitions.

All metadata describes an object. That object may be a physical thing (e.g., book, automobile, person) or other data (e.g., file, data element, database). Cataloging metadata is a set of descriptive properties of the object that are intended to facilitate the search/discovery of objects and to track/manage the objects. Cataloging metadata is used in registries as the official record of the object in the registry. Cards in a library card catalog contain metadata about books that both describe the book and provide pointers for finding the actual book in the library.

Administrative metadata is a set of descriptive properties that are intended to enable the monitoring and control of the object. Security attributes are metadata used by service and transmission software to control access to data; geotagging and creation timestamp attributes

are metadata in the images produced by digital cameras, for example, that describe the conditions under which the image was created and are captured by monitoring the behavior of the camera.

Structural/semantic metadata is a set of descriptive properties or specifications that overtly specify the structure or describe the meaning of data. XML Schemas are metadata that prescribe and are used to validate XML documents.

Metadata plays a crucial role in the development and use of information systems. Metadata enables information systems to capture information about what is “in” the information system and make “smarter,” more adaptable decisions, such as in securely controlling access to classified data. Metadata is also particularly important for search and discovery of items within and across Army networks; metadata recorded in registries provides a concise, fast, and known search space for consumers looking for particular kinds of information or artifacts.

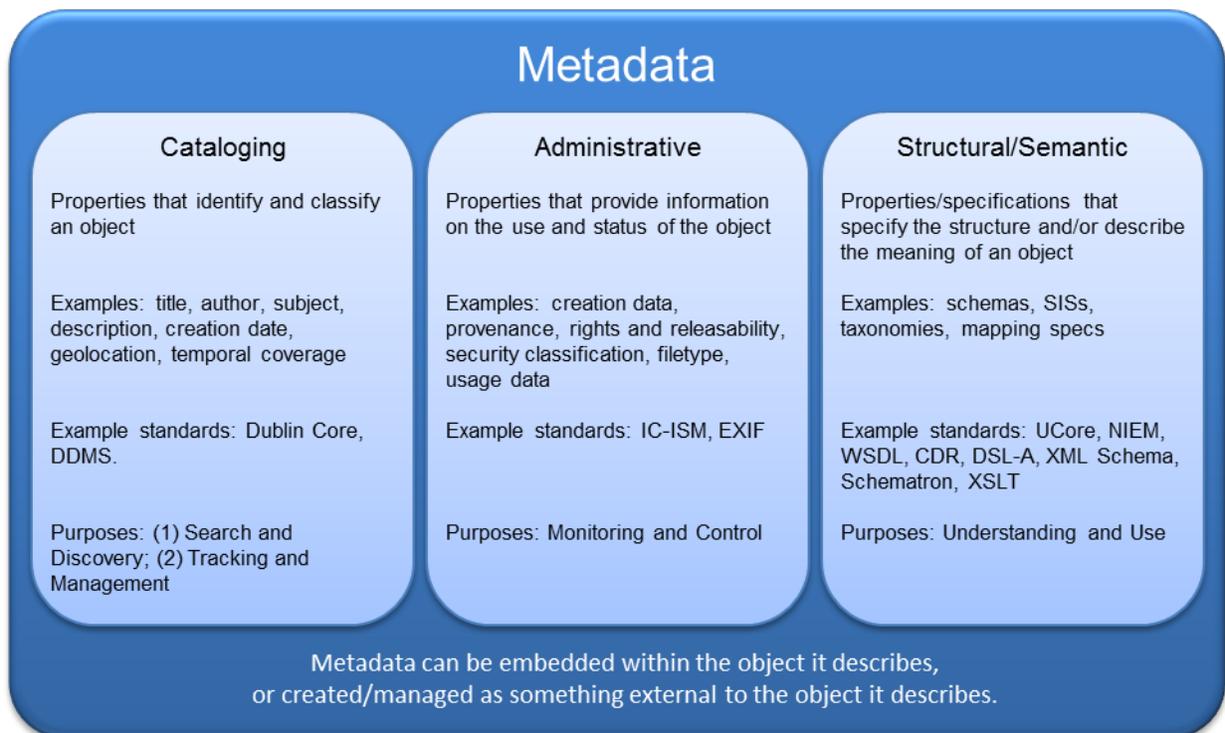


Figure 10: Metadata Categories

Principle DSD-24: Metadata is important in enterprise information systems for three (3) reasons: (1) search, discovery and understanding of enterprise assets; (2) monitoring and control of enterprise assets; and (3) adaptive, real-time operational control of information system behavior.

Business Rule DSD-24a: Metadata management strategies and standards should be adopted and incorporated into information systems designs as described in *ADF: Metadata Management* [51].

Business Rule DSD-24b: Metadata data models (i.e., the schemas that define metadata) should be designed with an anticipation of reuse by other organizations.

Business Rules DSD-24c: A DDMS metacard should be associated with an exchanged data asset, file or message [35].

The search and discovery usage of metadata meets the DoD information sharing goal of visibility and is implemented, in part, in DoD and Army registries.

DIL/LCE	Applications/devices in LCEs may use metadata and should not be responsible for metadata management. Metadata use, however, will play an important role in assessing the timeliness and relevance of data sent to and received from LCEs in DIL environments.
----------------	---

7.6.2 Registration

The DoD and Army's vision for net-centric operations is in large part about *unanticipated* information sharing (i.e., unanticipated reuse of data). The intent is for *all* Army information to be available to *anyone* (with the proper authorization) throughout the Army.

Unanticipated information sharing implies that the seeker of information may not know where to find the information needed and that data providers don't know who may need to use their data. Data providers must, therefore, make their information visible and discoverable (i.e., easy to find) and accessible (i.e., easy to obtain) to be of service to unanticipated users. The first step in attaining these goals is the *registration* of the provider's data asset with recognized registries and repositories in the DoD and Army. Registration applies to both data assets and data services. The importance of registration leads to the following principle and business rules.

Principle DSD-25: Registration of services, data, and metadata in recognized, authoritative DoD and Army registries make services and data visible and discoverable.

Business Rule DSD-25a: Fielded data services shall be registered with the DSE 2.0.

Business Rule DSD-25b: IESs, schemas, data models, service WSDLs, and other metadata shall be registered with the DoD DSE 2.0.

Business Rule DSD-25c: Data services under development should be registered with the DSE 2.0.

Business Rule DSD-25d: Data standards specifications (e.g., IESs, data services) that have proven useful by demonstration of successful and widespread adoption should be registered with the DoD Technology Standards and Profile Registry (DISR) [59].

Business Rule DSD-25e: Registries should monitor data discovery actions and feed statistics back to PMs, CIO/G-6, and the ADB to improve AIA and development processes.

Business Rule DSD-25f: Local registries should not be established, but in cases where the development of a local registry is justified, the local registry should be federated with DoD and Army level registries when functional overlaps exist.

Business Rule DSD-04a: If a data asset contains information that may support: (1) an Business Enterprise Architecture (BEA) end-to-end process [37] or (2) a JCA capability [31], then the data asset should be submitted to and registered with the DoD DSE 2.0 for consideration and certification as an Authoritative Data Source. The submission of a data asset for consideration may entail adjudication of competing claims of authority or jurisdiction.

Principle DSD-26: The use of “tags” to describe register-able items enables and facilitates the discovery of those items by search engines.

Business Rule DSD-26a: The DDMS [35] should be used to “tag” items submitted to registries.

Registration of services and the registration and submission of artifacts can be accomplished directly through the DSE 2.0 [25]. The process is described in *ADF: Metadata Management* [51]; this document also provides a list of recognized, authoritative registries in the DoD and Army.

The DSE 2.0 is shown as a registry on the data asset layer of the End-State Information Sharing Framework illustrated in Figure 7.

The DSRA Information Architecture [18] defines a Registration pattern that describes the data/service registration process.

DII/LCE	Applications/devices in LCEs may use registries, but most registration guidance is not applicable to LCEs. For example, data services in LCEs need not be registered.
----------------	---

8. Data Delivery and Use (DDU)

8.1 Enabling Data Delivery and Use

While data exchange is primarily about interoperability among information systems, information sharing is ultimately about providing information to people executing business processes, e.g., a depot item manager looking for a replacement part or a soldier receiving enemy location information. The delivery and use of data by human users through information system UI in pursuit of their missions should be traceable and managed all the way from the data asset from which it originated to the UI.

This section covers three aspects of data delivery and use:

- Information Requirements Traceability: the connections between data assets through UI display to the use of information in mission processes;
- Dashboards and Portals: data aggregation and data access UI paradigms for delivery and search/access of data pertinent to mission processes; and
- Business Intelligence (BI): the use of “big data” to support analytics, prediction, and insight.

8.2 Information Requirements Traceability

Information requirements are the descriptions of the information needed to drive enterprise processes and capabilities. Information is data that is interpreted and used (or created and stored) within the context of business or mission processes. In order to ensure that the right information is available and can be supplied to the right end-users in the Army and that it meets their information requirements, the use of data should be traceable from its storage location to the points where it is used or created within the Army.

DoDAF defines views that support the documentation of these traces. DoDAF OV-2s, OV-3s, and OV-5bs illustrate the flow of information between operational activities along need-lines.

Principle DDU-01: The information required to execute mission area processes, and to enable collaboration (i.e., information sharing) between Business/Mission Area processes, is supplied by (and is traceable to) specific sources of data.

Business Rule DDU-01a: A mapping (or trace) from a data asset to the business or mission area processes supported by the data asset should be developed and maintained. The mapping/trace may be documented using the DoDAF models identified in Table 4.

The DoD BEA [37] identifies fifteen (15) end-to-end business processes that may serve as a basis for traceability.

This is a long-term, systems engineering objective that is a guiding beacon in the context of near term Army information sharing objectives.

DIL/LCE

This guidance is important to LCEs in DIL environments because it is critical that warfighters (at the tactical edge) get the correct, latest, and highest quality data. The ability to trace the lineage of data enables the validation of the correctness, timeliness, and quality of the data received by warfighters.

8.3 Dashboards and Portals

Dashboards and portals are styles of UI design for providing visibility of and access to aggregated data. A dashboard is a visual display of data that is pulled from multiple sources with an objective to provide a particular set of information to a particular user for a particular purpose so the information can be analyzed and monitored at a glance. For example, a commander may use a dashboard to see soldier and enemy positions, status of soldier support resources, and weather conditions simultaneously and thereby obtain situational awareness; a fleet manager may use a dashboard to see the status and readiness of the fleet.

A portal is an integrated, centralized, personalized, web-based user interface and single-point-of-access to information sources, applications and collaborative services. Army Knowledge Online (AKO) is an excellent example of a portal.

Figure 11 illustrates differences between dashboards and portals. Dashboards “pull” data from sources to present to the user, while portals provide the users with access paths to data.

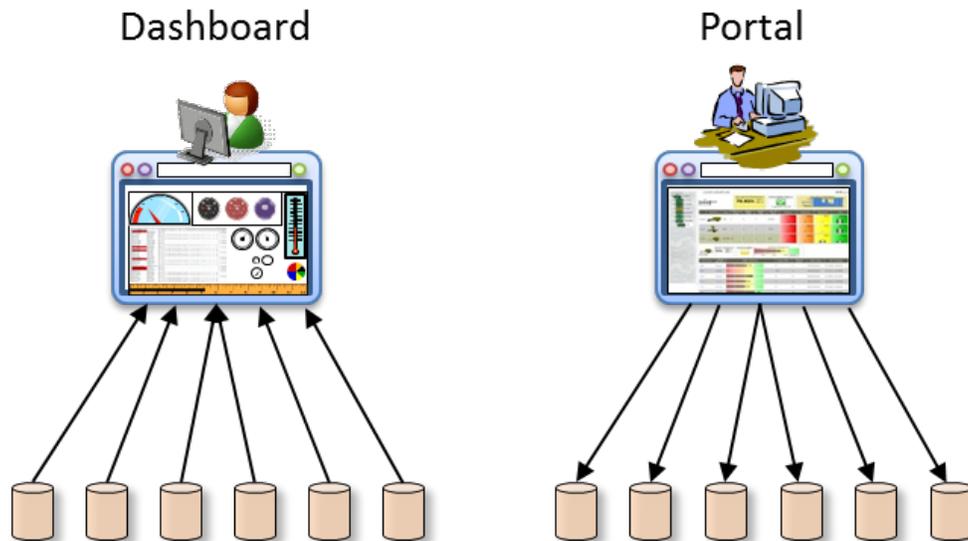


Figure 11: Dashboard and Portals

Principle DDU-02: Consistent UI design and deployment facilitates and promotes information sharing by providing a uniform and understood visual display for accessing and receiving information.

Business Rule DDU-02a: Dashboard and Portal UI design should adopt and follow the recommendations provided in *ADF: Dashboards and Portals* [54].

DIL/LCE	<p>Business Rule DDU-02b: In general, UI design for applications/devices in LCEs should be as simple and uncluttered as possible.</p> <p>If possible, users from tactical environments should be included in UI design working groups to ensure their requirements are included and met.</p>
----------------	--

8.4 Business Intelligence

"Business Intelligence" (BI) refers to the technologies, practices, and processes associated with the aggregation of data (usually from many different sources), the analysis of that data, and the delivery of information generated from that analysis to leaders, managers, and operators within a business. The primary purpose of BI is to enable and drive more effective, well-informed, and timely business decisions and improve business performance and competitiveness. BI is an application layer process within the End-State Information Sharing Framework illustrated in Figure 6; it is a user/consumer of the data assets and data service provided on the Army networks.

The kinds of data assets and services needed to support BI will vary greatly depending on the kind of analysis undertaken or information sought. The information sharing goals of visibility, accessibility, and understandability and the guidance provided by the AIA support to meet those goals provides a very basic foundation for general BI. Data asset/service requirements for analytic BI applications may require a data warehouse or a cloud-based solution for handling very large volumes of data input to the analysis. Many BI solutions assume a data warehouse as an underlying source of data for analysis.

Principle DDU-03: BI provides valuable business performance and competitive information to leaders, managers, and operators. Different kinds of BI require support by data asset and data service solutions that may be unique to the kind of BI.

Business Rule DDU-03a: BI solutions should adopt and follow the recommendations provided in *ADF: Business Intelligence (BI) Description* [55].

9. Secured Availability (SA)

9.1 Enabling Secured Availability and Access

The connectivity offered by the GIG and envisioned by the DoD information sharing objectives enables the greater information availability, access, and dissemination needed to meet the DoD information sharing objectives. This increase in information access capability also brings greater risk from threats that seek to obtain the information or disrupt its flow. These risks can only be countered with systemic security mechanisms that are not only woven into, but are an integral part of system design, development, fielding, and operation. PMs, PoRs, FDMs, and other roles throughout the Army must simultaneously meet the responsibility to share information with the Army while at the same time protect information against the risk of compromise. SA involves protecting the confidentiality, integrity, and availability (CIA) of Army information. The following items are aspects of CIA, and follow the guidelines in the DoD IEA:

- **Identity.** Security begins with identity. Establishing the clear and unambiguous identity of an individual user, software agent, or device (all referred to as “principals”) is the cornerstone of security.
- **Authentication.** Authentication is the act of verifying the identity or other attributes of a principal requesting access to protected data or a data service resource, based on a set of claims about that identity or attributes and the evidence given to support those claims. It is also the process of verifying the source and integrity of data. (Adapted from [54].)
- **Authorization.** Authorization is the “access privileges granted to a user, program, or process, or the act of granting those privileges.” [54] The security system must determine if the principal is authorized to perform the requested operation on a protected resource. This is accomplished by determining if the attributes of the principal, the request, and the current environment are sufficient to meet the policies protecting of the resource. For example, a given role or security clearance may be required for access (principal attribute), the resource might only be accessible if the request was encrypted (request attribute), or the resource might only be accessible during business hours (environment attribute).
- **Transmission Level Security.** Involves “measures (security controls) applied to transmissions in order to prevent interception, disruption of reception, communications deception, and/or derivation of intelligence by analysis of transmission characteristics such as signal parameters or message externals.” [54] Examples of transmission level security include Secure Sockets Layer (SSL) and Transport Layer Security (TLS).
- **Message Level Security.** Securely getting a message from one principal to another involves Integrity and Non-repudiation:
 - **Integrity.** Security mechanisms must ensure that messages and other data have not been modified in transit. A comparison means must be in place so that the data transmitted can be verified to match that received.
 - **Non-repudiation.** The “assurance that the sender of the information is provided with proof of delivery and the recipient is provided with proof of the sender's identity. Provides protection against an individual falsely denying having performed a particular action.”
- **Security markings.** Data - particularly data in transit from one point to another – must be appropriately and securely marked with security tags (e.g., metadata such as IC ISM [33]) that indicate its security level.

Ensuring the security of data and data services during maintenance and operation also involves:

- **Threat Assessment and Vulnerability Testing.** New threats emerge over time; security requires continuous monitoring⁵.
- **Security Management.** Changing mission needs and threats means that the access privileges and resource allocation will need to change quickly and easily based on policy and security attributes.
- **Security Mechanisms Management.** The security infrastructure must be planned and designed to ease the implementation of new security technologies. The security infrastructure must also be modular and separated from the main business logic of the resources being secured. Such a modular infrastructure is thus reusable across multiple resources.
- **Auditing and Logging.** The security system must also provide management personnel with a means of identifying what threats have attempted to access protected systems. Both successful and failed attempts at access should be logged and time tagged.

The principles and business rules presented in the following sections focus narrowly on data security and data service access security; they present only a partial picture of the security requirements for Army systems. They do not comprise a complete description of security guidance or requirements. AR 25-2 Information Assurance [14] is the broad and general starting point for data, information, and access security for the Army.

9.2 Information Assurance (IA)

Information Assurance (IA) (see AR 25-2 [14]) is the broad security activity charged with protecting the GIG and DoD systems from threats. IA involves the development and fielding of security mechanisms and defenses, incorporation of those mechanisms/defenses into information system development, assessing the security posture of a system, and monitoring, detecting, and responding to security incidents. In addition to AR 25-2, the DoD IA policies are currently governed by the DoD Information Assurance and Certification and Accreditation Process (DIACAP) [57], where the Army CIO/G-6 issues authorizations to operate for systems to be certified for use. Furthermore, the Army NETCOM provides a Certificate of Networkiness (CoN) process to evaluate applications, systems, and services for usage on Army networks and insures all security and maintainability of those systems are appropriately taken into account. Implementers of Army information systems, applications, services, and systems leverage the Defense Information Systems Agency (DISA) Security Technical Implementation Guides (STIGs) [37] that provide guidance for developers and checklists for security auditors. SA is the part of IA that focuses on protecting data while still making it readily available to authorized consumers.

The following principles are general IA requirements that apply to all Army information system design and development. While they encompass data and data service security, they address a scope broader than that of this document. They are included here for comprehensiveness.

Principle SA-01: A comprehensive, thorough, and conscientiously-implemented IA program will ensure, to the highest degree possible, the protection and security of Army information systems and the information they contain and process.

⁵ “Continuous monitoring” is the practice of having the controls and processes in place to always have an up-to-date security posture and obviate the need for regular security testing. See definition in Appendix B.2.

Business Rule SA-01a: To protect and secure Army information and data, information system design and development shall comply with the requirements stipulated in AR 25-2, Information Assurance [14].

Compliance with AR 25-2 encompasses the DIACAP, CIO/G-6 Authority to Operate (ATO), NETCOM CoN, and DISA STIGs.

Security awareness does not stop once a service or data asset is fielded. Security awareness, monitoring, and maintenance are important throughout the operational lifecycle of the service or data asset.

Principle SA-02: Risk assessment is an essential component of protecting information and data assets.

Business Rule SA-02a: Information systems and data assets should routinely be subjected to continuous monitoring throughout the information system lifecycle IAW AR 25-2 [14] Section 7-1.

While AR 25-2 is comprehensive in coverage of information security requirements, the following sections amplify or extend security requirements that pertain to information sharing, data exchange, and data services.

9.3 Data Security

The following sections address data security exclusive of data access. SA starts at the “ground floor” of data security.

9.3.1 Classification

Information security in the Army is governed by AR380-5, *Department of the Army Information Security Program* [15]. A security classification is applied to a piece of information, e.g., “range of missile XYZ.” Any subsequent expression “...incorporating, restating, paraphrasing, or generating in new form...” ([15] clause 2-1) of that information retains that classification, and is called a *derivative classification*. The security classification of data is thus derived from the security class of the information it represents.

Information security entails data security. The security requirements and mechanisms that are applicable to data depend on the implementation characteristics of the data. Security requirements and mechanisms for data-at-rest, for example, are different than those for data-in-transit.

Principle SA-03: All data have a security or protection level.

Business Rule SA-03a: All exchanged data in XML format (i.e., “data-in-transit”) shall include security level markings specified IAW the IC ISM metadata standard [33]. Markings shall include all pertinent classification data, such as declassification date.

Business Rule SA-03b: All persistent, stored data (i.e., “data-at-rest”) shall, directly or indirectly, be marked with a security classification level.

An example of indirect assignment of a security classification is that the security classification level of a system can be assumed to apply also to the data within that system.

Business Rule SA-03c: Unstructured data and files containing office-work products (e.g., documents, spreadsheets, and presentations), shall be marked with security markings IAW AR 380-5 [15]; if no guidance is provided in the regulation that is specific to the kind of file, it shall be marked as if it were a physical paper product.

Principle SA-04: The security classification of data is the same as (i.e., derived from) that of the information expressed by, represented by, or contained in the data.

Business Rule SA-04a: The security classification level of data shall be assigned by the data owner per AR 25-2 4-6.c IAW with a Security Classification Guide (SCG; see [15] Chapter 2, Section IV), such as the *Army Tactical Information Systems Security Classification Guide* [59].

A significant risk concerning the security classification of data is the possibility that disparate pieces of unclassified data may become classified when combined. There is no concrete guidance within the Army or DoD for recognizing or preventing this kind of situation or occurrence other than raising awareness of the possibility that it may occur.

9.3.2 Media and Devices

Data does not exist except as encoded on some physical media or device, e.g., bit patterns on magnetic drive, signals transmitted over a network. Data security entails the physical security of data encoding medium. Requirements for the physical security of media and devices is outside the scope of this document but is addressed in AR 25-2 [14].

DIL/LCE	Any “data-at-rest” that is created or collected by devices in LCEs should not be persistently and solely stored on the device. Data should be cached and during any periods of unprioritized connectivity uploaded to more capable CEs, e.g., Command Post and Data Center CEs.
----------------	---

9.3.3 Encryption

Data must be protected from unauthorized disclosure whether at rest or in transit. This is accomplished by encrypting data in a manner deemed acceptable by the Army for the level of classification and the medium on/in which the data is encoded, e.g., magnetic media, network packet transmission, solid state drives.

Principle SA-05: Data confidentiality policies and controls respect, enforce, and implement the privacy rights of information owners by protecting sensitive information from unauthorized access.

Business Rule SA-05a: Data shall be encrypted based on the security classification of the data and medium on/in which it is encoded. For the Army, data shall be encrypted IAW AR 25-2 [14].

9.3.4 Transfer

Data transfer is the movement (or, more precisely, the copying) of data between one physical location and another (i.e., from one physical medium to another), including movement via transmission over a network (i.e., data-in-transit).

Principle SA-06: Data that is transferred between locations is subject to threats while in transit and can be satisfactorily protected with security controls appropriate to the method/mechanism of transfer.

Business Rule SA-02a: Information systems and data assets should routinely be subjected to continuous monitoring throughout the information system lifecycle IAW AR 25-2 [14] Section 7-1.

Business Rule SA-06a: Data-in-transit shall be encrypted based on the security classification of the transmitted data IAW AR 25-2 [14].

Business Rule SA-06b: Data-in-transit shall be signed using National Security Agency (NSA)-approved signature algorithms (see [60]).

A *security domain* is a collection of people, networked hardware, systems, and data over which one or more security policies apply. The collection may, but need not, be a network that is physically disjoint from other networks, e.g., SIPRNET, NIPRNET. A *Cross Domain Solution* (CDS) is security solution explicitly developed to enable data to be exchanged across security domains. Data transfer across security domains is a particularly important security consideration.

Business Rule SA-06c: Data that is transferred between security domains via physical transport of physical media (e.g., “sneaker-net”) shall be evaluated and sanitized IAW AR 25-2 [14] clause 4-16 and Information Assurance Best Business Practice *Data Transfer Across Security Domains* [60].

Business Rule SA-06d: Data shall not be electronically transmitted (“data-in-transit”) between security domains unless the CDS over which the data was transmitted was developed and approved IAW AR 25-2 [14] clause 4-21 and the Information Assurance Best Practice *Cross Domain Solutions* [62].

Business Rule SA-06e: The security markings of data transferred between security domains shall be updated for the new security domain as appropriate.

9.3.5 Disposal

When data has reached the end of its useful life it needs to be disposed of in an appropriate manner. This will ensure that the data does not become compromised and minimizes the impact of maintaining and processing irrelevant or out-of-date information on Army systems.

Principle SA-07: Neglected, abandoned, or “residual” data is a security risk.

Business Rule SA-07a: Data that is moved, no longer useful, left on discarded devices, should be destroyed, wiped, purged, or sanitized IAW AR 25-2 [14] Section 4-20.

Disposal is particularly important in cloud-based data assets when terminating the consumer - cloud provider relationship. See *Termination Migration Plan* in Appendix G.4.2.

9.4 Data Service Security

Data service security is a subset of general service security. Data service security should be consistently and uniformly implemented across the Army through the development and use of data service security guidance, which is an intersecting subset of data service guidance (see Section 7.4.2) and IA (see 9.2) access controls.

Principle SA-08: The use of data service security guidance will make the incorporation of security controls into Army data services easier, improve the security of Army data, and reduce the costs of security solutions.

Business Rule SA-08a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data service security guidance. The same body, or an allied body, should develop, test, and promote service security mechanisms, tools, and resources. The data service security guidance shall be derived from, complementary to, and consistent with Information Assurance access controls. The guidance shall include:

- Security Information and Event Management (SIEM) policies/technology to improve data security by building-in the monitoring capabilities for threat prevention and auditing capabilities for intrusion forensics;

Business Rule SA-08b: Data service developers shall adopt and use data service security guidance in the design, development, fielding, and operation of data services. In particular, data service developers should adopt and use the following:

- WS-Security is a standard extension to SOAP to apply security to web services. The protocol specifies how integrity and confidentiality can be enforced on messages and allows the communication of various security token formats, such as SAML, Kerberos, and X.509. It uses XML Signature and XML Encryption to provide end-to-end security.
- Security Assertion Markup Language (SAML) [63] is a standard for the exchange of principal authentication and attribute information between clients, services, and security services.
- eXtensible Access Control Markup Language (XACML) [64] is a standard that defines a declarative access control policy language and processing model to evaluate authorization requests according to the rules defined in policies.
- WS-SecurityPolicy [64] is a standard that can be used to enhance WSDL specifications to represent and exchange security policy information.
- Tactical Service Security System (TS3) security handlers [66] is a software product developed by the Army for incorporation into web services development and deployment to support the consistent implementation of security mechanisms. TS3 incorporates WS-Security and SAML.
- Identity and Access Management (IdAM) - Reference Architecture [67] specifies requirements for the implementation of Identity Management for authentication of principles and controlling access to data services.

DIL/LCE

These data service security mechanisms may not be applicable/suitable in LCEs due to size/performance overhead. Alternative and equivalent-or-stronger service security mechanism may be used in LCEs in place of the mechanisms listed above.

A uniform and consistent Army-wide approach to authorizing data service consumers is the key first step to Secured Availability.

Principle SA-09: Authorization of principals to access data services via uniform access control mechanisms across the Army will ease availability of and access to data while maintaining a high level of security.

Business Rule SA-09a: Data stewards or an ADB designee shall establish, publish, and promote authoritative access control policies and mechanisms for data service access, including:

- Managing access by Attribute-Based Access Control (ABAC) policy to control the level of access by both anticipated and unanticipated users;

- Developing and fielding an enterprise-level Policy Decision Service that is used by web services to authorize and authenticate access; and
- Establishing an access control policy for each protected resource.

Business Rule SA-09b: Army security access control policies and mechanisms shall be used in the design, development, fielding, and operation of data services. In particular, data services shall comply with the access control requirements stipulated in AR 25-2 [14] Section 4-12.

Business Rule SA-09c: Security elements of DSRA [18] and the ADF-Data-related Security chapter [14] should be applied in planning and implementation of security measures.

Business Rule SA-09d: Data service design, development, and fielding shall follow NCES's security policies [60] [69].

Business Rule SA-09e: Data service design, development, and fielding shall follow the guidelines defined in the DISA STIGs [37].

Business Rule SA-09f: Data services shall obtain a CoN issued by NETCOM before fielding.

Authorization of a consumer to use a data service does not necessarily mean that the consumer can access all data available through via the service. If the data available from a service falls within several classifications, a consumer may be entitled to some, but not all, of the data.

DIL/LCE	Authorization of principals in LCEs requesting access to enterprise-level services should follow these business rules. Authorization of principals to access services in a LCE should take place locally within the LCE and should not use an enterprise-level authorization service. In all cases, authorization/authentication should not impede priority operations in LCEs.
----------------	---

Principle SA-10: Authorization to use a data service is distinct from authorization to access the data provided by the service. Security and protection of data is ensured by matching the authorization level of an authenticated identity with the security level of the data.

Business Rule SA-10a: A data service shall authorize data access by validating the security level (i.e., privileges) of the authenticated identity of a requesting consumer against the security level of the requested data.

NOTE: If the security level of the data available in a service is not distinct from the security or protection level of the service itself, the security level of the data shall be assumed to be the same as the security level of the service.

Business Rule SA-10b: A data service should validate the "need-to-know" of the consumer requesting the data. Data service consumers and data service providers should adopt the *XML Data Encoding Specification for Need-To-Know Metadata* [70] for automating the validation of need-to-know.

Business rules SA-10c: A data service should exchange data IAW Intelligence Community Multi-Audience Tearline (IC-MAT) [71] when appropriate based on the content of the exchange.

The need-to-know validation addresses requirement at AR 25-2 4-6.f(10).

10. Governance⁶

Loosely-coupled and dynamic data service capabilities offer significant and measurable benefits, as discussed in the Army Net-Centric Data Strategy [31]. If not properly governed, however, the enterprise IT development can quickly evolve into a disarrayed collection of unmanageable services that are not interoperable. Successful information sharing in a distributed environment requires the enforcement of policies, common standards and schemas, and guidelines that promote a consistent approach across the enterprise ensuring common understanding and interoperability.

Data governance refers to the overall management of the availability, usability, integrity and security of the data employed in an enterprise. A sound data governance program includes a governing body or council, a defined set of procedures, and a plan to execute those procedures. Governance involves policy making, decision arbitration, executive sponsorship, and day-to-day operational administration. The AIA is a principal instrument of Army Data Governance.

The objective of data governance within the Army is to enable the Army to make more effective and efficient use of data assets in achieving operational goals. The Army has established the position of Chief Data Officer (CDO) to shepherd data management practices and lead the Army's data governance process.

10.1 Chief Data Officer (CDO)

The CDO is responsible for developing, implementing, and enforcing data standards (Army, DoD, and Federal) and the Army's data enterprise strategy. The CDO will develop, approve, and certify data management processes and products. The desired end state is a data governance framework that results in a data service environment that allows warfighters and decision-makers access to information in a timely and secure manner regardless of their environment.

The CDO's role includes leveraging Reference Architectures to identify standards in the acquisition lifecycle phase to facilitate data interoperability. Additionally, the CDO will approve and certify data management processes and products.

The ADB and the ANCDs CoE have been created to provide support to the CDO. This support forum will provide the CDO with the ability to develop Army data governance positions, policy, and technology for use both internally within the Army and externally within the DoD, Combatant Commands, Sister Services, and Industry.

⁶ The material presented in this section was adapted from architecture.army.mil and www.milsuite.mil and the Army Data Board charter [75].

10.2 Army Data Board (ADB)

The ADB is the senior directive body for development of coordinated Army enterprise positions on data strategy, standards and execution in conformance with the Army COE. This board serves as the senior adjudication body for Army enterprise data issues; acts as the final authority across the Army enterprise for data standards, policies and practices; coordinates data-sharing efforts across the Army enterprise; serves as a certification/waiver approval authority for targeted standards as delegated by the CDO; and collects and disseminates best practices and lessons learned for the data community.

Chaired by the Army CDO, the board is comprised of Army Data Stewards. Data Stewards are 1-2 Star level General Officer/Senior Executive Service (SES) equivalent nominated by the Assistant Secretaries of the Army, Deputy Chiefs of Staff, Army Commands, and other areas as defined by the Army CDO.

The ADB is supported by an operational body, the ADC, that is responsible for executing the policies and directives issues by the ADB. This body serves as the ADB's initial adjudication body and its membership consists of FDMs from each of the Army's Assistant Secretaries of the Army, Deputy Chiefs of Staff, Army Service Commands and Direct Reporting Units or other areas as determined by individual Data Stewards and the CDO.

Appendix A References

- [1] Department of Defense Chief Information Officer. *Department of Defense Information Enterprise Architecture*, Version 1.2. 7 May 2010.
[document] http://jitc.fhu.disa.mil/jitc_dri/pdfs/dod_ia_v1_2_7_may_2010.pdf
[web site] <http://dodcio.defense.gov/Home/Initiatives/DIEA.aspx>
- [2] Department of Defense Chief Information Officer. *Department of Defense Information Enterprise Architecture (DoD IEA) Version 2.0 Volume I – Management Overview of the DoD IEA*. July 2012.
http://dodcio.defense.gov/Portals/0/Documents/DIEA/DoD%20IEA%20v2.0_Volume%20I_Description%20Document_Final_20120730.pdf
- [3] Department of Defense Chief Information Officer. *Department of Defense Information Enterprise Architecture (DoD IEA) Version 2.0 Volume II – IEA Description*. July 2012.
http://dodcio.defense.gov/Portals/0/Documents/DIEA/DoD%20IEA%20v2%200_Volume%20II_Description%20Document_Final_20120806.pdf
- [4] U.S. Army CIO/G-6. *Common Operating Environment Architecture, Appendix C to Guidance for 'End State' Army Enterprise Network Architecture*, 1 Oct 2010.
[CAC required] <https://www.us.army.mil/suite/doc/38201362>
- [5] McHugh, John M, Secretary of the Army. *Army Directive 2009-03, Army Data Management*. 30 Oct 2009.
http://www.apd.army.mil/pdf/files/ad2009_03.pdf
- [6] Office of the Assistant Secretary of Defense, Networks and Information Integration (OASD/NII). *Reference Architecture Description*. June 2010.
[web site] <http://dodcio.defense.gov/Home/Initiatives/DIEA.aspx>
[document] http://dodcio.defense.gov/Portals/0/Documents/DIEA/Ref_Archi_Description_Final_v1_18Jun10.pdf
- [7] Department of Defense Chief Information Officer. *Data Sharing in a Net-Centric Department of Defense*. DoD Directive 8320.02. 2 Dec 2004, certified current 23 Apr 2007.
<http://www.dtic.mil/whs/directives/corres/pdf/832002p.pdf>
- [8] ASA(ALT) Common Operating Environment Implementation Plan, Draft, Nov 2011 v3.0
<http://www.bctmod.army.mil/downloads/pdf/COE.pdf>
- [9] CIO/G-6. *Guidance for LWN 2020 and Beyond End State Architecture*. Version 0.8, 11 March 2013.
- [10] US Army Office of Business Transformation. *2013 Annual Report on Business Transformation Providing Readiness at Best Value*. 1 Mar 2013.
<http://dcmo.defense.gov/publications/documents/2013%20Army%20Report%20on%20Business%20Transformation%20March%202013.pdf>
- [11] Chairman of the Joint Chiefs of Staff. *Interoperability and Supportability of Information Technology and National Security Systems*. Instruction CJCSI 6212.01F, 21 Mar 2012.
http://www.dtic.mil/cjcs_directives/cdata/unlimit/6212_01.pdf
- [12] Headquarters, Department of the Army. *Army Knowledge Management and Information Technology*, AR 25-1. 4 Dec 2008.
http://armypubs.army.mil/epubs/pdf/r25_1.pdf

-
- [13] Headquarters, Department of the Army. *Information Technology and Support Services*, DA PAM 25-1-1. 25 Oct 2006.
http://armypubs.army.mil/epubs/pdf/p25_1_1.pdf
- [14] Headquarters, Department of the Army. Information Assurance, AR 25-2. DRAFT Feb 2012.
- [15] *Information Security Program*, AR380-5, Department of the Army. 29 Sep 2000.
http://armypubs.army.mil/epubs/pdf/R380_5.pdf
- [16] *US Army Enterprise, Army Data Framework (ADF): Overview*, Version 1.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 23 April 2012.
[CAC required] <https://www.intelink.gov/go/y4iqGSc>
- [17] *US Army Enterprise Data Strategy Reference Architectures Part 1: Overview, Common Components, and Common Vocabulary*, Version 1.7, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 30 Jul 2010.
[CAC required] <https://www.intelink.gov/go/b1T0jLU>
- [18] *Data Services Layer – Army Service Interface Specifications Overview*, Version 1.2.1, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 30 Jun 2010.
[CAC required] <https://www.intelink.gov/go/la06GL8>
- [19] Intelligence Community and Department of Defense Content Discovery & Retrieval Integrated Project Team. *IC/DoD Content Discovery & Retrieval Reference Architecture*. V1.1, 25 Feb 2011.
[CAC required] <https://www.intelink.gov/sites/odni/cio/i2e/focus/iads/cdript/default.aspx>
[CAC required] <https://www.intelink.gov/wiki/CDRIPT>
- [20] Intelligence Community and Department of Defense Content Discovery & Retrieval Integrated Project Team. *IC/DoD Content Discovery & Retrieval Specification Framework*. V1.0 Draft, 9 May 2011.
[CAC required] <https://www.intelink.gov/sites/odni/cio/i2e/focus/iads/cdript/default.aspx>
[CAC required] <https://www.intelink.gov/wiki/CDRIPT>
- [21] Namespace Management for the Army Enterprise – General Solution, Version 1.0, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 29 Sep 2009.
[CAC required] <https://www.intelink.gov/go/OD7he4p>
- [22] Namespace Management for the Army Enterprise – XML Namespaces, Version 1.0, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 29 Sep 2009.
[CAC required] <https://www.intelink.gov/go/KZnpoqp>
- [23] *Rules for Cross-Cutting Capability (CCC) Information Exchange Specifications (IES) in Interface Specifications*, Version 1.3, 28 Dec 2013.
<https://www.intelink.gov/go/adlzEja>
- [24] DoD Architecture Framework,
<http://dodcio.defense.gov/dodaf20.aspx>
- [25] AIA Compliance Assessment: Compliance Matrix and Assessee Briefing Deck
[CAC Required] <https://www.intelink.gov/go/zbGhwAe>
-

-
- [26] Erl, Thomas. *SOA Principles*.
<http://www.soaprinciples.com/>
- [27] Data Services Environment (DSE)
[CAC required] <https://metadata.ces.mil/dse/homepage.htm>
- [28] *Department of Defense (DoD) Information Technology Standards Program (ITSP) Management Plan*. Defense Information Systems Agency (DISA). 19 January 2007.
- [29] *US Army Enterprise, Army Data Framework (ADF): Data Quality Management*, Version 1.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 15 Jun 2012.
[CAC required] <https://www.intelink.gov/go/q9mZPaf>
- [30] *US Army Enterprise, Army Data Framework (ADF): Data Warehouse Description*, Version 1.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 22 Apr 2009.
[CAC required] <https://www.intelink.gov/go/T11NyK6>
- [31] *Joint Capability Areas*
[CAC required] https://intellipedia.intelink.gov/wiki/Joint_Capability_Areas
- [32] Department of Defense Chief Information Officer. *Net-Centric Data Strategy*, 9 May 2003.
<http://dodcio.defense.gov/Portals/0/documents/Net-Centric-Data-Strategy-2003-05-092.pdf>
- [33] *Department of Defense Interface Standard – Variable Message Format (VMF)* MIL-STD-6017, 19 Jul 2006.
- [34] *XML Data Encoding Specification for Information Security Marking (ISM)*, Version 9. Office of the Director of National Intelligence. 17 July 2012.
[Web site] <http://www.dni.gov/index.php/about/organization/chief-information-officer/information-security-marking-metadata>
[Document] <http://www.dni.gov/files/documents/CIO/ICEA/ISMPublic.zip>
- [35] *Department of Defense Discovery Metadata Specification (DDMS)*, Verion 4.0.1. Defense Information System Agency (DISA), Program Executive Officer, GIG Enterprise Services. 11 Nov 2011.
[Web site] <http://metadata.ces.mil/dse/irs/DDMS/>
- [36] *US Army Enterprise Data Strategy Reference Architectures Part 3: Information Architecture*, Version 1.5, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 22 Feb 2010.
[CAC required] <https://www.intelink.gov/go/zaCWWYC>
- [37] DoD Office of the Deputy Chief Management Officer. *Business Enterprise Architecture (BEA) 10.0*. 14 Feb 2013
<http://dcmo.defense.gov/products-and-services/business-enterprise-architecture/10.0/classic/index.htm>
<http://dcmo.defense.gov/products-and-services/business-enterprise-architecture/10.0/classic/html/end2end.htm>
- [38] Wikipedia. *Watson (computer)*.
[http://en.wikipedia.org/wiki/Watson_\(computer\)](http://en.wikipedia.org/wiki/Watson_(computer))

-
- [39] *US Army Enterprise, Army Data Framework (ADF): Data Aspects of Cloud Computing*, Version 1.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 15 Mar 2012.
[CAC required] <https://www.intelink.gov/go/tlrpDIm>
- [40] Wikipedia. *Big data*.
http://en.wikipedia.org/wiki/Big_data
- [41] *The NIST Definition of Cloud Computing* (SP 800-145). U.S. Department of Commerce, NIST, Computer Security Division, Information Technology Laboratory. Peter Mell & Timothy Grance. Sep 2011.
<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- [42] Kundra, Vivek, U.S. Chief Information Officer. *Federal Cloud Computing Strategy*, 8 Feb 2011.
http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf
- [43] *Challenging Security Requirements for US Government Cloud Computing Adoption*. Cloud Computing Program, Information Technology Laboratory, National Institute of Standards and Technology (NIST). NIST Working Draft. Nov 2011
http://www.nist.gov/manuscript-publication-search.cfm?pub_id=912695
- [44] Federal Risk and Authorization Management Program (FedRAMP)
<https://cio.gov/cyber-security-2/fedramp/>
- [45] *US Army Enterprise, Army Data Framework (ADF): Enterprise Resource Planning*, Version 1.1. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 31 May 2012.
[CAC required] <https://www.intelink.gov/go/CcF84j5>
- [46] World Wide Web Consortium, *SOAP*, Version 1.2.
<http://www.w3.org/TR/soap/>
- [47] *Representational state transfer*
http://en.wikipedia.org/wiki/Representational_state_transfer
- [48] *Army Data Services Layer Developers' Guide*, Version 1.2.1, Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 30 Jul 2010.
[CAC required] <https://www.intelink.gov/inteldocs/view.php?fDocumentId=345438>
- [49] Common Data Services Framework.
[CAC required] <https://www.us.army.mil/suite/page/647548>
- [50] *US Army Enterprise, Army Data Framework (ADF): Master Data Management*, Version 1.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 15 July 2010.
[CAC required] <https://www.intelink.gov/go/JsoBPxW>
- [51] Department of Defense. Under Secretary of Defense for Acquisition, Technology, and Logistics. *Unique Identification (UID) Standards for a Net-Centric Department of Defense*. DoD Directive 8320.03. 23 Mar 2007.
www.dtic.mil/whs/directives/corres/pdf/832003p.pdf
- [52] *US Army Enterprise, Army Data Framework (ADF): Metadata Management*, Version 0.6. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 29

- July 2010.
[CAC required] <https://www.intelink.gov/go/r5DOZzr>
- [53] DoD Information Technology Standards and Profile Registry (DISR)
[CAC required] <https://acc.dau.mil/CommunityBrowser.aspx?id=148577>
- [54] *US Army Enterprise, Army Data Framework (ADF): Dashboards and Portals*, Version 2.0. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 03 May 2012.
[CAC required] <https://www.intelink.gov/go/yhXQIXf>
- [55] *US Army Enterprise, Army Data Framework (ADF): Business Intelligence (BI) Description*, Version 1.1. Office of the Army Chief Information Officer, Army Net-Centric Data Strategy Branch, 20 July 2010.
[CAC required] <https://www.intelink.gov/go/E72IICd>
- [56] Committee on National Security Systems. *National Information Assurance (IA) Glossary*. CNSS Instruction No. 4009, 26 April 2010
www.cnss.gov/Assets/pdf/cnssi_4009.pdf
- [57] DoD Information Assurance Certification and Accreditation Process (DIACAP) -- Online DIACAP Training - DIACAP Overview
<http://iase.disa.mil/diacap/>
- [58] *Application Security and Development Security Technical Implementation Guide (STIG)*, Version 3, Release 2. Defense Information Systems Agency (DISA), 24 Jul 2008,
http://iase.disa.mil/stigs/downloads/zip/u_application_security_and_development_stig_v3r2_20101029.zip
- [59] *Army Tactical Automated Information Systems Classification Guide*, Army CIO/G-6, Network Enterprise Technology Command (NETCOM), Enterprise Systems Technology Activity (ESTA), Office of Information Assurance and Compliance (OIA&C). Apr 2008.
- [60] Committee on National Security Systems. *National Information Assurance Policy on the Use of Public Standards for the Secure Sharing of Information Among National Security Systems*. CNSS Policy No. 15, 01 Oct 2012.
http://www.cnss.gov/Assets/pdf/CNSSP_No%2015_minorUpdate1_Oct12012.pdf
- [61] *Data Transfer Across Security Domains*, Version 1. 03-EC-T-0002. 23 May 2006.
https://www.milsuite.mil/book/servlet/JiveServlet/download/14373-2-58403/03ECT0002_May232006.pdf
- [62] *Army Cross Domain Solution Procedures*, Version 2.0. 07-EB-O-0001. US Army CIO/G-6 Cyber Directorate. Sep 2011.
https://www.milsuite.mil/book/servlet/JiveServlet/download/14378-7-74047/07-EB-O-0001_CrossDomainSolutions_14%20SEP11.pdf
- [63] OASIS. Security Assertion Markup Language (SAML). March 2002
http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security
<http://saml.xml.org/saml-specifications>
- [64] Organization for the Advancement of Structured Information Standards (OASIS). *eXtensible Access Control Markup Language (XACML)*, Version 3.0, 22 January 2013.
http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml

-
- [65] OASIS, WS-SecurityPolicy 1.2, 1 Jul 2007
<http://docs.oasis-open.org/ws-sx/ws-securitypolicy/200702/ws-securitypolicy-1.2-spec-os.html>
- [66] *Developers Guide for the Tactical Services Security System (TS3)*, Version 11.12.30.0. C-E LCMC Software Engineering Center. 06 Jan 2012.
[CAC required] <https://www.us.army.mil/suite/kc/13671973>
- [67] *U.S. Army - Identity and Access Management (IdAM) - Reference Architecture (RA) v1.0*, Draft Version 0.97, 04 June 2012.
- [68] Defense Information Systems Agency (DISA). *NCES Service Security Design Specification for Service Security*, Version 0.4.6, 13 Aug 2007.
<https://www.us.army.mil/suite/doc/9830907>
- [69] Defense Information Systems Agency (DISA). *NCES Service Security Interface Specification for Service Security*, Version 0.4.6, 13 Aug 2007.
<https://www.us.army.mil/suite/doc/9830780>
- [70] *XML Data Encoding Specification for Need-To-Know Metadata*, Version 7. Office of the Director of National Intelligence. 17 July 2012.
<http://www.dni.gov/index.php/about/organization/chief-information-officer/need-to-know-metadata>
- [71] *XML Data Encoding Specification for Multi-Audience Tearlines (IC-MAT)*, DNI CIO, Version 7.0, 17 July, 2012
[Web site] <http://www.dni.gov/index.php/about/organization/chief-information-officer/multi-audience-tearline>
[Document] <http://www.dni.gov/files/documents/CIO/ICEA/MATPublic-Light.zip>
- [72] Blohm, Gary W., Army Chief Data Officer. *Army Data Board Charter*. Version 5.02. 24 Jan 2013.
<http://architecture.army.mil/Governance/ADBoard.html>
- [73] FGM, Inc. *Data Services Environment 2.0 Concept of Operations*. 20 Jul 2012.
https://metadata.ces.mil/dse-hlp/documents/DSE_CONOPS.pdf
- [74] Office of the Director of National Intelligence. *Intelligence Community Directive Number 503*. 15 Sep 2008.
http://www.dni.gov/files/documents/ICD/ICD_503.pdf

Appendix B Acronyms and Definitions

B.1 Acronyms and Abbreviations

ABAC	Attribute-Based Access Control
ABCD	Army Bulk CBM+ Data
ADB	Army Data Board
ADC	Army Data Council
ADF	Army Data Framework
ADF	Artifact Data Framework
ADS	Authoritative Data Source
AI	Artificial Intelligence
AIA	Army Information Architecture
AKO	Army Knowledge Online
AMC	Army Materiel Command
ANCDS	Army Net-Centric Data Strategy
ANSI	American National Standards Institute
API	Application Program Interface
AR	Army Regulation
ASA(ALT)	Assistant Secretary of the Army for Acquisition, Logistics, and Technology
ASCII	American Standard Code for Information Interchange
ATO	Authority To Operate
BI	Business Intelligence
BML	Battle Management Language
BTP	Business Transformation Plan
C&A	Certification and Accreditation
C2 Core	Command and Control Core
CAC	Common Access Card
CBM	Conditions-Based Maintenance
CBRN	Chemical, Biological, Radiological, Nuclear
CCC	Cross-Cutting Capability
CCE	Cloud Computing Environment
CDF	Common Data Format
CDO	Chief Data Officer
CDR	Content Discovery and Retrieval
CDS	Cross Domain Solution
CDSF	Common Data Services Framework

CE	Computing Environment
CECOM	Communication Electronics Command
CIA	confidentiality, integrity and availability
CIO	Chief Information Officer
CIR	Computing Infrastructure Readiness
CJSC	Chairman of the Joint Chiefs of Staff
CNSS	Committee on National Security Systems
CoE	Center of Excellence
COE	Common Operating Environment
COI	Community of Interest
CoN	Certificate of Networthiness
CONOPS	Concept of Operations
CONUS	Continental United States
COTS	Commercial-Off-The-Shelf
CR	Communications Readiness
CRIS	Common Relational Information Schema
CRUD	Create Read Update Delete
CV	Capability View
DA	Department of the Army
DaaS	Data as a Service
DADM	Data Asset Development and Management
DAMA	Data Management Association
DCGS-A	Distributed Common Ground Systems - Army
DDF	Data Description Framework
DDL	Data Description Language
DDMS	DoD Discovery Metadata Specification
DDU	Data Delivery and Use
DIACAP	DoD Information Assurance and Certification and Accreditation Process
DIL	Disconnected, Intermittent, Limited
DISA	Defense Information Systems Agency
DISR	DoD Information Technology Standards and Profile Registry
DIV	Data and Information View (DoDAF)
DMBOK	Data Management Body of Knowledge
DoD	Department of Defense
DoDAF	DoD Architecture Framework
DQM	Data Quality Management

DRM	Data Reference Model
DSC	DCGS-A SIPR Cloud System
DSD	Data and Services Deployment
DSE	Data Services Environment
DSL-A	Data Services Layer – Army
DSPI	Data Storage, Processing, and Integration
DSRA	Data Strategy Reference Architecture
EDF	Enterprise File Delivery
ERP	Enterprise Resource Planning
ESB	Enterprise Service Bus
ESG	Executive Steering Group
ESM	Enterprise Service Management
ESTA	Enterprise Systems Technology Activity
ETL	Extract, Transform, and Load
FDM	Functional Data Manager
FIPS	Federal Information Processing Standards
FISMA	Federal Information Security Management Act
FORSCOM	U.S. Forces Command
GA	General Army
GCDS	GIG Content Delivery Services
GEIA	Government Electronics & Information Technology Association
GFIEDM	Global Force Management Information Exchange Data Model
GIG	Global Information Grid
GML	Geography Markup Language
HBSS	Host-Based Security System
HQDA	Headquarters, Department of the Army
IA	Information Assurance
IAW	In Accordance With
IC	Intelligence Community
ICD	Intelligence Community Directive
IE	Information Enterprise
IEA	Information Enterprise Architecture
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IES	Information Exchange Specification
IESS	Information Exchange Specification Standards

IETF	Internet Engineering Task Force
ISM	Information Security Markings
ISO	International Standards Organization
ISR	Intelligence, Surveillance and Reconnaissance
IT	Information Technology
ITSP	Information Technology Standards Program
ITU	International Telecommunications Union
JCA	Joint Capability Area
JDBC	Java Database Connectivity
KML	Keyhole Markup Language
LCE	Limited Computing Environment
LCMC	Life Cycle Management Command
MAT	Multi-Audience Tearline
MDF	Model Description Framework
MDM	Master Data Management
MIL-STDs	Military Standards
MIME	Multipurpose Internet Mail Extensions
MIMOSA	Machinery Information Management Open Systems Alliance
NATO	North Atlantic Treaty Organization
NETCOM	Network Enterprise Technology Command
NIEM	National Information Exchange Model
NIPR	Non-classified IP Router Network
NIST	National Institute of Standards and Technology
NMES	Namespace Management Enterprise Solution
NOA	NetOps Agility
NSA	National Security Agency
NSG	National System for Geospatial-Intelligence
OASIS	Organization for the Advancement of Structured Information Standards
OBT	Office of Business Transformation
ODBC	Open Database Connectivity
OGC	Open Geospatial Consortium
OIA&C	Office of Information Assurance and Compliance
OSA	Open System Architecture
OV	Operational View
PAM	Pamphlet
PEO	Program Executive Office

PLCS	Product Life-Cycle Support
PM	Project/Program Manager
POA&M	Plan of Action and Milestones
POC	Point of Contact
PoR	Program of Record
RA	Reference Architecture
REST	Representational State Transfer
SA	Secured Availability
SAML	Security Assertion Markup Language
SCG	Security Classification Guide
SCRUD	Search-Create-Read-Update-Delete
SEC	Software Engineering Center
SES	Senior Executive Service
SIEM	Security Information and Event Management
SIPR	Secure Internet Protocol Router
SIS	Service Interface Specification
SLA	Service Legal Agreement
SME	Subject Matter Expert
SOA	Service-Oriented Architecture
SOAP	not an acronym; originally stood for Simple Object Access Protocol
SSL	Secure Sockets Layer
STANAG	Standardized Agreement
STEP	Standard for the Exchange of Product Model Data
STIG	Security Technical Implementation Guide
SvcV	Service View
SV	System View
TDQM	Total Data Quality Management
TLS	Transport Level Security
TRADOC	Training and Doctrine Command
TRM	Technical Reference Model
TS3	Tactical Service Security System
UDDI	Universal Description, Discovery and Integration
UI	User Interface
UID	Universal Identifier
URI	Uniform Resource Identifier
USB	Universal Serial Bus

USMTF	United States Message Text Format
VAUS	Visible, Accessible, Understandable, Secure
VMF	Variable Message Format
W3C	World Wide Web Consortium
WBS	Work Breakdown Structure
WSDL	Web Services Description Language
XACML	Extensible Access Control Markup Language
XSLT	Extensible Stylesheet Language Transform
XML	Extensible Markup Language
XSD	XML Schema Description language

B.2 Definitions

Anticipated Information Sharing

The act of sharing information in which the consumer and provider has a prior arrangement for sharing information.

Business Rule

A recommendation, requirement, directive, stipulation, or imperative that asserts what should/shall be done to meet or implement a principle. A business rule may be an end-state objective.

Continuous Monitoring

“The process implemented to maintain a current security status for one or more information systems or for the entire suite of information systems on which the operational mission of the enterprise depends. The process includes: 1) The development of a strategy to regularly evaluate selected IA controls/metrics, 2) Recording and evaluating IA relevant events and the effectiveness of the enterprise in dealing with those events, 3) Recording changes to IA controls, or changes that affect IA risks, and 4) Publishing the current security status to enable information sharing decisions involving the enterprise.” [54]

Data Asset

“Any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a web site that returns data in response to specific queries (e.g., www.weather.com) would be a data asset. A human, system, or application may create a data asset.” (From DoDD 8320.02 [7].)

Data Asset, Structured

A bounded, finite (though possibly dynamically changing and/or distributed) collection of data that meets the following conditions:

- is governed by a single schema (i.e., physical data model), or set of integrated schemas (e.g., nested schemas), which entails that all data type names for the data elements are unique;

- the boundaries of the container define a managed identifier space within which all data element identifiers (e.g., relational data keys, record IDs, or XML “id” attributes) are unique; and
- has a single unique, holistic identifier (e.g., path/file name, database id, message id).

Most often, “data asset” refers to a database, but it may be used to refer to anything meeting the stated criteria (e.g., a message).

Data Exchange

The physical and mechanical transference of data (via copying or transmission) from one location to another without consideration of meaning or intent of the data.

Data Integration

The process of combining data from two or more sources and producing a single unified, consistent, and cohesive view of the combined data. Generally, the objective is to produce a set of data that represents the same information that is represented by the input data sets, though this need not always be the case.

Data-based Integration

A data-centric strategy, approach, or architecture that (1) is designed to enable or implement an integrated, comprehensive, consistent, enterprise-spanning data deployment and management solution, and that (2) enables enterprise application interoperability. Another, more accurate term is “Data-based System Integration,” i.e., (system) integration that is based on data.

Data Transformation

The process of converting data from one physical format to another.

Data Translation

Data Transformation that preserves (to the greatest extent possible) the information content of data input to the transformation process in the resultant output data.

Information Exchange Specification (IES)

A set of specifications that govern the physical data format for data exchanged among members of a community. The specifications must include schemas, definitions of schema elements & relationships, and specifications of constraints. The specifications may include other relevant or supporting content that augments the primary specification content.

Information Requirement

The information needed to drive or execute an enterprise business process or make a decision.

Information Sharing

The exchange of data with the purpose of conveying, sharing, or communicating information pertinent to some purpose or process.

Information System (IS)

(From AR 25-2 [14]) “...a set of information resources organized for the collection, storage, processing, maintenance, use, sharing, dissemination, disposition, display, or transmission of information. As part of the set of information resources, an IS includes its own operating system(s), firmware, hardware (or all of the above) to support a single mission or across a range of missions.”

Information Technology Governance/Guidance Documentation

The body of documentation⁽¹⁾ that governs or guides the design, development, deployment, operation, and retirement of information systems⁽²⁾.

- (1) policies, directives, guidelines, architectures, laws, templates, procedures, regulations, methods, patterns, practices, standards, specifications, tenets, principles, doctrine, rules, instructions, examples, training material.
- (2) IT in general, systems, hardware, software, infrastructure, applications, services, architectures, databases.

Mediation

The process of translating data between a source format and a receiving format using a neutral, “third party” format (e.g., an IES) that mediates the exchange. Mediation involves two (2) data translation steps: from source format to mediating format, and from mediating format to target format.

Principle

A generalized statement of position that is accepted as true or valid, and often reflects values, beliefs, or convictions on the “right” or “best” way to do or achieve a result or fulfill a mission. Principles guide decision-making and actions; a principle is not an end-state objective.

Semi-structured Data

Data that is not governed by a schema and that has (1) an informal, internal structure, and/or (2) internal markings or tags, that enable software applications to extract, “understand” and process “fine-grained semantics” of the data at same level as structured data.

Service

A software component/application that performs a defined function and is accessible/callable by consumer agents via a published interface/API on/over a network.

Service Interface Specification (SIS)

The formal specification of the functionality of a service and the API for a service.

Structured Data

Data that is governed by a schema and can be validated against that schema; structured data typically represents “fine-grained semantics” where a data element name indicates the meaning of a data value at a level that can be used by software applications.

Unanticipated Information Sharing

The act of searching for and obtaining information from a provider in which there is no *a priori* arrangement between the consumer and the provider to obtain/provider the information. The consumer and provider may have no prior knowledge of one another.

Unanticipated User

A service consumer whose information requirements (i.e., “data needs”) were not explicitly considered in the design, development, and fielding of a service.

Unstructured Data

Digital representation of text, video, audio, or imagery that is intended for interpretation by human agents. The text is typically natural language expressions that would be used in email, social networking sites, news sites, etc.

Appendix C List of Principles and Business Rules

The following list of principles and business rules is provided for summary and convenience.

Label	Principle/Business Rule	Page
GA-01	Principle GA-01: Data is an Enterprise Asset. Information is Enterprise Currency. Knowledge is an Enterprise Resource.	21
GA-02	Principle GA-02: Data is a physical representation of information but is not the same thing as information.	21
GA-03	Principle GA-03: Effective decision-making and effective process execution in the Army requires effective Information Sharing.	22
GA-04	Principle GA-04: Information creators and managers have a responsibility and obligation to make their data visible and accessible to consumers throughout the Army.	22
GA-04a	Business Rule GA-04a: Information creators and managers shall have a plan and schedule (i.e., implementation plan) for making their data available to the Army (if not already available).	22
GA-05	Principle GA-05: The information that drives decision-making and Army processes is available to authorized consumers regardless of their location or the time of their request.	22
GA-05a	Business Rule GA-05a: Data should be accessible by authorized consumers across the Army within the security restrictions on the data.	22
GA-06	Principle GA-06: Compliance with Army governance and guidance documentation will enable, facilitate, and promote effective information sharing among Army information systems and meet DoD information sharing objectives.	23
GA-06a	Business Rule GA-06a: Data stewards or an ADB designee shall develop, maintain, and promote data, service, and architecture governance and guidance documentation and shall assess compliance to the documentation.	23
GA-06b	Business Rule GA-06b: Architects and developers shall ensure that systems comply with the following data governance and architectural guidance documentation, as applicable: <ul style="list-style-type: none"> • Army Knowledge Management and Information Technology, AR 25-1 [12]; • The COE Architecture [4]; • The Army Data Framework (ADF) [14]; • The Data Strategy Reference Architecture [14]; • Content Discovery & Retrieval (CDR) [19]; and • Data Services Layer - Army [18]. 	23
GA-06c	Business Rule GA-06c: Data stewards or an ADB designee should develop/acquire, test, and promote tools and resources that support adherence to or compliance with data, service, and architecture governance and guidance documentation.	24
GA-06d	Business Rule GA-06d: Architects and developers shall adopt, implement, or use standards and governance/guidance documentation in the preferential order presented in Table 3 (adapted from [57]).	24
GA-06e	Business Rule GA-06e: Army governance and guidance shall take into account constraints of LCEs and DIL environments.	24

Label	Principle/Business Rule	Page
GA-06f	Business Rule GA-06f: If the governance and guidance documentation does not explicitly address LCEs or DIL environments, LCEs should comply with the documentation cited in GA-06b as applicable to/within the constraints of the LCE and potential DIL environments.	24
GA-07	Principle GA-07: The effectiveness of Army governance documentation can be measured (in part) by the cost savings that results from adopting the guidance/solutions.	24
GA-07a	Business Rule GA-07a: PoRs/PMs/Data Stewards should separately track costs of development, deployment, and sustainment of enterprise data, data services, and COI activities, to measure, manage, and improve efficiency and effectiveness of AIA.	24
DADM-01	Principle DADM-01: Information of value to the Army is represented by structured, semi-structured, and unstructured data.	27
DADM-02	Principle DADM-02: Effective information sharing is based on clear, unambiguous, and consistent management of structured data. Physical data models (aka schemas) are a necessary mechanism for managing the format and semantics of (i.e., the information conveyed by) structured data.	27
DADM-02a	Business Rule DADM-02a: A schema (i.e., physical data model) shall be developed and maintained for each structured data asset (e.g., database, data service interface, or message format). In DoDAF architectures, the schema would be a DIV-3 Physical Data Model.	28
DADM-02b	Business Rule DADM-02b: An inventory (list) of the data assets within the scope of responsibility of a program or system shall be developed and maintained.	28
DADM-03	Principle DADM-03: The use of data model guidance will improve interoperability, information sharing, and increase the value of data produced by, and available to, the Army by engendering common perspectives, technologies, and approaches in the data model development process.	28
DADM-03a	Business Rule DADM-03a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data model guidance.	29
DADM-03b	Business Rule DADM-03b: Army data model guidance should be used in the analysis, development, and implementation of data assets.	29
DADM-03c	Business Rule DADM-03c: Data should be logically separated (i.e., decoupled) from applications by applying design and analysis guidance provided in the data model guidance.	29
DADM-04	Principle DADM-04: Data models that are designed in accordance with Army data modelling guidelines and principles will increase the longevity, usefulness, and reusability of the data model, and will make information sharing (in both the near and long-term) easier and more effective.	29
DADM-04a	Business Rule DADM-04a: Data stewards or an ADB designee shall develop, promote and maintain data modelling guidelines and principles as of part of the data model guidance.	29
DADM-04b	Business Rule DADM-04b: Data models should be developed in accordance with Army data modelling guidelines and principles.	29

Label	Principle/Business Rule	Page
DADM-04c	Business Rule DADM-04c: Data model (schema) design, specification, development, and fielding shall adhere to the Army Namespace Management Enterprise Solution (NMES) [21] [22].	29
DADM-04d	Business Rule DADM-04d: Data models (schemas) documentation should include a very clear definition of the scope of the information represented by the data model and the mission use of that information.	29
DADM-04e	Business Rule DADM-04e: Data models (schemas) should be designed with an anticipation of unanticipated users and future scope changes that are needed to address changing mission requirements.	29
DADM-04f	Business Rule DADM-04f: Data modelling guidance for data stored in LCEs and moved into/out of DIL environment should be tailored to the limitations of the devices in the LCE and to the limitations of the DIL environment.	29
DADM-05	Principle DADM-05: Effective decisions require high-quality data.	30
DADM-05a	Business Rule DADM-05a: Data stewards or an ADB designee shall develop, promote, and maintain a Data Quality Management (DQM) program.	30
DADM-05b	Business Rule DADM-05b: DQM processes, programs, or standards should be adopted and applied in data system design, development, and operation.	30
DADM-06	Principle DADM-06: A comprehensive DQM program will produce and ensure high-quality data.	30
DADM-06a	Business Rule DADM-06a: A DQM program should follow or adopt the DoD Guidelines for Total Data Quality Management (TDQM) as outlined in the Army Data Framework (ADF): Data Quality Management (DQM) [28] Section 2.1.1.	30
DADM-06b	Business Rule DADM-06b: A TDQM program should establish and use a standard set of data quality dimensions to evaluate and measure data quality as outlined in the ADF-DQM [28] Section 2.1.2.	30
DADM-06c	Business Rule DADM-06c: A TDQM program should adopt and implement the DQM best practices outlined in the ADF-DQM [28] Section 2.4.	31
DADM-06d	Business Rule DADM-06d: Data quality assurance tools, mechanisms, and practices should be incorporated into system architectural specifications to ensure the quality of input and generated data and, thus, prevent low-quality data from even getting into the system. See ADF-DQM [28] Section 3.1.	31
DADM-06e	Business Rule DADM-06e: A TDQM program should implement governance procedures that clearly define the roles and responsibilities for DQM as outlined in ADF-DQM [28] Section 3.2.	31
DADM-07	Principle DADM-07: A clear, robust, and well-defined data integration process is critical to ensuring data quality when data is imported into a local data asset from multiple external data assets.	32
DADM-07a	Business Rule DADM-07a: Data assets/systems should adopt the data integration process outlined in the ETL process description in the ADF: Data Warehouse [19].	32
DSD-01	Principle DSD-01: The need for Information Sharing may be anticipated or unanticipated.	35

Label	Principle/Business Rule	Page
DSD-01a	Business Rule DSD-01a: Army data governance and architectural guidance documentation shall include strategies for addressing anticipated and unanticipated information sharing.	35
DSD-02	Principle DSD-02: Data models that are formalized and adopted as IESs will facilitate and enable effective information sharing within a community.	36
DSD-02a	Business Rule DSD-02a: IESs shall be formally documented in accordance with Army policies, templates, and other requirements governing IESs. At a minimum, the IES shall include a schema, the definitions of schema elements and the relationships among them, and the definition of any extra-schema constraints governing the validity of data that conforms to the schema.	36
DSD-02b	Business Rule DSD-02b: The data models upon which IESs are based shall follow, adhere to, or comply with Army Data Model Guidance (see Section 6.3).	36
DSD-02c	Business Rule DSD-02c: The data models (schemas) in IESs intended for use in DIL environments shall include a mapping/conversion to a compressed, binary physical exchange format unless a network impact study is conducted and a waiver is obtained.	37
DSD-03	Principle DSD-03: Reusing published IESs facilitates interoperability across the Army.	37
DSD-03a	Business Rule DSD-03a: Communities should pursue the adoption of an IES in the following preferential order: <ul style="list-style-type: none"> • Adopt and use a published IES or Information Exchange Standard (see Appendix E.4) as-is; • Research metadata/schema repositories such as the DoD DSE 2.0 for a data model(s) relevant to the community's information sharing requirements and adopt that data model(s); • Modify and adopt a published IES or data model discovered in a metadata repository; and • Develop and adopt a community-specific IES. 	37
DSD-04	Principle DSD-04: Timely, effective, and accurate decision-making depends on timely and accurate information; timely and accurate information depends on the availability and quality of Authoritative Data.	38
DSD-04a	Business Rule DSD-04a: If a data asset contains information that may support: (1) an Business Enterprise Architecture (BEA) end-to-end process [37] or (2) a JCA capability [31], then the data asset should be submitted to and registered with the DoD DSE 2.0 for consideration and certification as an Authoritative Data Source. The submission of a data asset for consideration may entail adjudication of competing claims of authority or jurisdiction.	38
DSD-04b	Business Rule DSD-04b: A data asset that has been certified as an ADS by the appropriate designated body shall be maintained in accordance with policies and procedures that govern ADSs.	38
DSD-04c	Business Rule DSD-04c: When timeliness/currency of data is important, real-time access to ADSs is preferred over non-ADS sources.	38
DSD-05	Principle DSD-05: Information (of a given type) that is available from a single source (rather than multiple sources) will reduce the possibility of conflicting information and increase the trustworthiness of the information.	38

Label	Principle/Business Rule	Page
DSD-05a	Business Rule DSD-05a: An ADB designee should analyze, evaluate, and plan the content of ADSs across the Army such that specific kinds of information (i.e., “data needs”) are not provided by multiple sources.	38
DSD-06	Principle DSD-06: Unstructured data is a valuable source of information.	39
DSD-06a	Business Rule DSD-06a: Data stewards or an ADB designee should (1) recommend technologies for the semantic analysis of unstructured data (e.g., written prose and recordings) and (2) develop an Army-wide strategy and guidance for harvesting and leveraging the information in unstructured data.	39
DSD-06b	Business Rule DSD-06b: A catalog of unstructured data assets (or locations) that contain information of value to the Army should be created and maintained. The DoD Discovery Metadata Specification (DDMS) [35] should be used to annotate/tag each unstructured data asset in the catalog to facilitate discovery.	39
DSD-06c	Business Rule DSD-06c: Unstructured data assets should be analyzed and tagged in accordance with Army strategic guidance.	39
DSD-07	Principle DSD-07: The use of cloud computing implementation guidance will improve interoperability, information sharing, and increase the value of data produced by, and available to, the Army by improving the reliability and availability of data.	40
DSD-07a	Business Rule DSD-07a: Data stewards or an ADB designee shall develop, publish, promote, and maintain cloud computing implementation guidance in accordance with Federal CIO and National Institute of Standards and Technology (NIST) guidance.	40
DSD-07b	Business Rule DSD-07b: Cloud computing implementation guidance shall be used in the design and development of cloud-based systems; in particular, the following guidance documentation shall be used: <ul style="list-style-type: none"> • Army Data Framework - Data Aspects of Cloud Computing [37]. 	40
DSD-07c	Business Rule DSD-07c: A migration/deployment plan shall be developed to guide the movement/deployment of data or data services to a CCE. The plan should follow, or be compatible with, the guidance provided in the Federal Cloud Computing Strategy [42] and appendix G.4. (See also Section 2.4.1 of Army Data Framework - Data Aspects of Cloud Computing [37].)	40
DSD-08	Principle DSD-08: The design decisions on the data storage implementation paradigm (e.g., relational, key-value, dimensional) and access methods (e.g., data services) for cloud-based data assets depend more on application/usage requirements than on Cloud Computing Technology.	41
DSD-08a	Business Rules DSD-08a: Cloud data storage and access should be designed to meet, and be suitable for, application usage requirements and leverage cloud computing benefits while also addressing cloud computing technology performance constraints and limitations.	41
DSD-09	Principle DSD-09: Data security and legal concerns are of greater significance when data is deployed in/to a Cloud Computing Environment when compared to conventional computing environments.	42

Label	Principle/Business Rule	Page
DSD-09a	Business Rule DSD-09a: A security plan shall be developed in conjunction with the movement/deployment of data or data services to a CCE. The security plan should address the security and privacy challenges presented in Challenging Security Requirements for US Government Cloud Computing Adoption [43] and the security/legal considerations presented in Appendix G.4.	42
DSD-09b	Business Rule DSD-09b: Data shall be deployed to a CCE In Accordance With (IAW) the security requirements and guidance provided by the Federal Risk and Authorization Management Program (FedRAMP) [42].	42
DSD-10	Principle DSD-10: A clear legal contract between a cloud service provider and cloud service consumer protects both the provider and consumer.	42
DSD-10a	Business Rule DSD-10a: A clear, unambiguous Service Level Agreement (SLA) or legal contract shall be prepared and signed by cloud service consumer and cloud service provider.	42
DSD-11	Principle DSD-11: ERP systems are the same as other software systems/applications in the Army when it comes to data exchange and information sharing: ERP systems interoperate with other Army systems and may serve as a data asset for consumers across the Army.	43
DSD-11a	Business Rule DSD-11a: ERP systems shall make their data available to consumers as appropriate to support interoperability with other systems, particularly Master or Authoritative data,	43
DSD-11b	Business Rule DSD-11b: ERP systems should be implemented IAW the guidance provided by the ADF: Enterprise Resource Planning [45].	43
DSD-12	Principle DSD-12: Data services contribute to meeting unanticipated information sharing requirements. Exposing data via data services makes data available to and accessible by unanticipated, authorized users.	43
DSD-12a	Business Rule DSD-12a: Data of value to the Army shall be made available to authorized consumers in the Army via data services.	43
DSD-13	Principle DSD-13: It is better (e.g., more cost effective) to reuse existing services than to develop a new service.	44
DSD-13a	Business Rule DSD-13a: Before the development of a data service is undertaken, service registries should be searched for both existing, fielded services, and services that are under development that could meet the requirements of the required data service. The DSE 2.0 is the primary service registry that should be researched.	44
DSD-13b	Business Rule DSD-13b: Where a service exists that meets the requirements of the required data service, the existing service shall be adopted and used and a new data service shall not be developed. If multiple services exist that fulfill a capability need, a data service shall be chosen and adopted in the priority order presented in Table 3.	44
DSD-13c	Business Rule DSD-13c: Where a data service (or services) exist that partially meet the requirements of a required data service, the owners of the required data service should engage the owners/maintainers of the existing data service to request a change in order to meet the requirements of the required data service. If a good-faith effort to change the existing data service fails in a reasonable length of time, then the required data service should be developed.	44

Label	Principle/Business Rule	Page
DSD-14	Principle DSD-14: If no service exists that fulfills a capability need, it is better (e.g., more cost effective) to use existing service interface specification standards for implementing the service than to implement a service with a unique, localized interface.	45
DSD-14a	Business Rule DSD-14a: If no web service exists that meets the requirements of the proposed data service, service and metadata repositories shall be searched for published/standardized data SIS that can fulfill the capability need. If no suitable SIS is found, a new SIS shall be developed and submitted to appropriate service and metadata repositories. If a suitable SIS is found, the specifications should be adopted and implemented as published; if the SIS only partially meets the requirements of the proposed data service, the authors of the SIS shall be engaged to request a change to the SIS; if a good-faith effort to change the SIS fails in a reasonable length of time, then a new SIS should be developed (or the partially suitable SIS extended) and submitted to appropriate service and metadata repositories.	45
DSD-15	Principle DSD-15: Some data services offer a capability that is best developed and provided by a single service that is used by consumers across the Army.	46
DSD-15a	Business Rule DSD-15a: Data stewards or an ADB designee shall be responsible for Army service portfolio management. This body (or bodies) shall research, identify, design, develop, deploy, and manage Enterprise Data Services that provide single-source capability to consumers across the Army.	46
DSD-15b	Business Rule DSD-15b: Data services should be considered for suitability as an Enterprise Data Service. If suitable, the data service should be brought to the attention of the Army's service portfolio management body.	46
DSD-16	Principle DSD-16: The use of data service guidance will ease the process and reduce the cost of both (1) creating and deploying data services, and (2) using data services.	47
DSD-16a	Business Rule DSD-16a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data service guidance. The same body, or an allied body, should develop, test, and promote service development support tools (e.g., CDSF) and resources.	47
DSD-16b	Business Rule DSD-16b: Army data service guidance should be applied in the design and development of data services.	47
DSD-16c	Business Rule DSD-16c: The design and development of a data service should: <ul style="list-style-type: none"> • follow the Army's data service development process; • adhere to the Army NMES [21] [22]; and • incorporate instruments or monitors to track performance and usage of the data service. 	47
DSD-16d	Business Rule DSD-16d: In DoDAF architectures, data services should be documented as a SvcV-4 Services Functionality Description view.	47
DSD-16e	Business Rule DSD-16e: Data service guidance shall take into account constraints of LCEs and DIL environments for the design of services to be deployed on LCE devices or in DIL environments.	47

Label	Principle/Business Rule	Page
DSD-17	Principle DSD-17: Data services that are designed and implemented in accordance with Army data service guidelines to be as “future proof” as possible will increase the longevity and reusability of the service.	47
DSD-17a	Business Rule DSD-17a: The design of a new data service should anticipate other potential users of the service (e.g., outside the known consumers of the service) and consider potential future uses. Data services should not be (effectively) a point-to-point service for a particular requirement. The data service should be generic and/or flexible so that it can be reused. This is what is meant by “loose coupling.”	47
DSD-17b	Business Rule DSD-17b: When a data service is changed, backward compatibility should be maintained so that existing clients of the service will not be affected.	47
DSD-17c	Business Rule DSD-17c: Data services should be designed to be scalable to handle more users than the number currently anticipated.	47
DSD-17d	Business Rule DSD-17d: The design of services to be deployed on LCE devices or in DIL environments may be exempted from general data service guidelines with a justification based on LCE or DIL environment constraints. Such services shall comply with any guidance specific to the LCE or DIL environment.	47
DSD-18	Principle DSD-18: Monitoring of the operation and performance of a data service will ensure that the service meets user expectations and serve as feedback to improve the data service design and development process.	47
DSD-18a	Business Rule DSD-18a: The performance of the data service should be monitored to ensure performance remains within acceptable limits to all users.	48
DSD-18b	Business Rule DSD-18b: PoRs should monitor data consumption actions and feed statistics back to PMs, CIO/G-6, and the ADB to improve AIA and development processes.	48
DSD-18c	Business Rule DSD-18c: The design of services to be deployed on LCE devices or in DIL environments may exclude monitoring capabilities with a justification based on LCE or DIL environment constraints.	48
DSD-19	Principle DSD-19: Master data is an important mechanism for integrating enterprise software systems. Master data management ensures that master data is always correct and current.	49
DSD-19a	Business Rule DSD-19a: Master data management processes should be included in interoperability architectures as described in ADF: Master Data Management [50].	49
DSD-19b	Business Rule DSD-19b: Master data management processes shall incorporate Unique Identifiers (UID) for significant Army assets as directed by DoD Directive 8320.03 Unique Identification (UID) Standards for a Net-Centric Department of Defense [51].	49
DSD-19c	Business Rule DSD-19c: Applications/services in LCEs may use master data but should not be responsible for any master data management functions.	49
DSD-20	Principle DSD-20: COIs are the basis for anticipated information sharing in the Army, and for defining and meeting interoperability requirements.	50

Label	Principle/Business Rule	Page
DSD-20a	Business Rule DSD-20a: Information producers and consumers that regularly share information should join (or form) and participate in COIs.	50
DSD-21	Principle DSD-21: Interoperability/collaboration is most effectively achieved, and an interoperability solution is most effectively designed and implemented, within a small community. Interoperability can be effectively scaled from small communities to larger communities.	50
DSD-21a	Business Rule DSD-21a: COIs should focus on data exchange among systems within a COI before considering data exchange within a broader community.	50
DSD-21b	Business Rule DSD-21b: COIs should have an Information SME who is responsible for knowing the data assets within the COI, the information they contain, and the relationships among them. This role is similar to that of an FDM with a scope of responsibility that covers the COI.	50
DSD-22	Principle DSD-22: Unambiguous interoperability between highly interactive applications (particularly those requiring data translation and data integration) requires overt, formal, and precise specification of data exchange pathways and information content of the exchanged data (i.e., the description of the anticipated information that is being shared).	50
DSD-22a	Business Rule DSD-22a: COIs shall create and maintain a DoDAF AV-2 Integrated Dictionary and should create and maintain a DIV-2 Logical Data Model.	50
DSD-22b	Business Rule DSD-22b: Applications/systems that interoperate frequently with other applications, particularly within a COI, should have an explicit documentation of the interoperations (e.g., resource flows or “data exchange pathways”) between the applications/systems. COIs should develop and document these interoperations using sets of DoDAF models [19]; the sets of models that document the interoperations are shown in Table 4; either or both sets (e.g., Business Process View or System Interaction View) should be developed.	51
DSD-22c	Business Rule DSD-22c: COIs should adopt, or develop and publish, one or more Information Exchange Specifications (IES) that represents the information available within the COI. COIs should adopt and use published IES standards where/when possible. The COI Information SME is responsible for the development of and is the custodian of the IES.	51
DSD-22d	Business Rule DSD-22d: The physical structure and information content of exchanged data shall be documented and governed by an IES.	51
DSD-23	Principle DSD-23: The formal specification of the relationship between two (2) different data models (i.e., schemas) or IESs is necessary to understand, monitor, and maintain consistent information content (i.e., semantics) of data during a data transformation process.	52
DSD-23a	Business Rule DSD-23a: Mapping specifications should be developed and published that define the relationship between local schemas and the IESs used to share information with collaborating partners. A “local schema” may be a schema published with a data service interface (i.e., “export schema” or “service schema”) or the schema of a data asset. The mapping specification shall be detailed enough to unambiguously define the data transformation process, and shall identify semantic gaps that result from the mapping.	52

Label	Principle/Business Rule	Page
DSD-23b	Business Rule DSD-23b: Formal mapping specifications shall be used to govern the transformation of data from a format conforming to one schema to format conforming to another schema.	52
DSD-24	Principle DSD-24: Metadata is important in enterprise information systems for three reasons: (1) search, discovery and understanding of enterprise assets; (2) monitoring and control of enterprise assets; and (3) adaptive, real-time operational control of information system behavior.	54
DSD-24a	Business Rule DSD-24a: Metadata management strategies and standards should be adopted and incorporated into information systems designs as described in ADF: Metadata Management [51].	54
DSD-24b	Business Rule DSD-24b: Metadata data models (i.e., the schemas that define metadata) should be designed with an anticipation of reuse by other organizations.	54
DSD-24c	Business Rules DSD-24c: A DDMS metacard should be associated with an exchanged data asset, file or message [35].	54
DSD-25	Principle DSD-25: Registration of services, data, and metadata in recognized, authoritative DoD and Army registries make services and data visible and discoverable.	55
DSD-25a	Business Rule DSD-25a: Fielded data services shall be registered with the DSE 2.0.	55
DSD-25b	Business Rule DSD-25b: IESs, schemas, data models, service WSDLs, and other metadata shall be registered with the DoD DSE 2.0.	55
DSD-25c	Business Rule DSD-25c: Data services under development should be registered with the DSE 2.0.	55
DSD-25d	Business Rule DSD-25d: Data standards specifications (e.g., IESs, data services) that have proven useful by demonstration of successful and widespread adoption should be registered with the DoD Technology Standards and Profile Registry (DISR) [59].	55
DSD-25e	Business Rule DSD-25e: Registries should monitor data discovery actions and feed statistics back to PMs, CIO/G-6, and the ADB to improve AIA and development processes.	55
DSD-25f	Business Rule DSD-25f: Local registries should not be established, but in cases where the development of a local registry is justified, the local registry should be federated with DoD and Army level registries when functional overlaps exist.	55
DSD-26	Principle DSD-26: The use of “tags” to describe register-able items enables and facilitates the discovery of those items by search engines.	56
DSD-26a	Business Rule DSD-26a: The DDMS [35] should be used to “tag” items submitted to registries.	56
DDU-01	Principle DDU-01: The information required to execute mission area processes, and to enable collaboration (i.e., information sharing) between Business/Mission Area processes, is supplied by (and is traceable to) specific sources of data.	57

Label	Principle/Business Rule	Page
DDU-01a	Business Rule DDU-01a: A mapping (or trace) from a data asset to the business or mission area processes supported by the data asset should be developed and maintained. The mapping/trace may be documented using the DoDAF models identified in Table 4.	57
DDU-02	Principle DDU-02: Consistent UI design and deployment facilitates and promotes information sharing by providing a uniform and understood visual display for accessing and receiving information.	58
DDU-02a	Business Rule DDU-02a: Dashboard and Portal UI design should adopt and follow the recommendations provided in ADF: Dashboards and Portals [54].	58
DDU-02b	Business Rule DDU-02b: In general, UI design for applications/devices in LCEs should be as simple and uncluttered as possible.	58
DDU-03	Principle DDU-03: BI provides valuable business performance and competitive information to leaders, managers, and operators. Different kinds of BI require support by data asset and data service solutions that may be unique to the kind of BI.	59
DDU-03a	Business Rule DDU-03a: BI solutions should adopt and follow the recommendations provided in ADF: Business Intelligence (BI) Description [55].	59
SA-01	Principle SA-01: A comprehensive, thorough, and conscientiously-implemented IA program will ensure, to the highest degree possible, the protection and security of Army information systems and the information they contain and process.	61
SA-01a	Business Rule SA-01a: To protect and secure Army information and data, information system design and development shall comply with the requirements stipulated in AR 25-2, Information Assurance [14].	62
SA-02	Principle SA-02: Risk assessment is an essential component of protecting information and data assets.	62
SA-02a	Business Rule SA-02a: Information systems and data assets should routinely be subjected to continuous monitoring throughout the information system lifecycle IAW AR 25-2 [14] Section 7-1.	62
SA-03	Principle SA-03: All data have a security or protection level.	62
SA-03a	Business Rule SA-03a: All exchanged data in XML format (i.e., “data-in-transit”) shall include security level markings specified IAW the IC ISM metadata standard [33]. Markings shall include all pertinent classification data, such as declassification date.	62
SA-03b	Business Rule SA-03b: All persistent, stored data (i.e., “data-at-rest”) shall, directly or indirectly, be marked with a security classification level.	62
SA-03c	Business Rule SA-03c: Unstructured data and files containing office-work products (e.g., documents, spreadsheets, and presentations), shall be marked with security markings IAW AR380-5 [15]; if no guidance is provided in the regulation that is specific to the kind of file, it shall be marked as if it were a physical paper product.	62
SA-04	Principle SA-04: The security classification of data is the same as (i.e., derived from) that of the information expressed by, represented by, or contained in the data.	63

Label	Principle/Business Rule	Page
SA-04a	Business Rule SA-04a: The security classification level of data shall be assigned by the data owner per AR 25-2 4-6.c IAW with a Security Classification Guide (SCG; see [15] Chapter 2, Section IV), such as the Army Tactical Information Systems Security Classification Guide [59].	63
SA-05	Principle SA-05: Data confidentiality policies and controls respect, enforce, and implement the privacy rights of information owners by protecting sensitive information from unauthorized access.	63
SA-05a	Business Rule SA-05a: Data shall be encrypted based on the security classification of the data and medium on/in which it is encoded. For the Army, data shall be encrypted IAW AR 25-2 [14].	63
SA-06	Principle SA-06: Data that is transferred between locations is subject to threats while in transit and can be satisfactorily protected with security controls appropriate to the method/mechanism of transfer.	63
SA-06a	Business Rule SA-06a: Data-in-transit shall be encrypted based on the security classification of the transmitted data IAW AR 25-2 [14].	64
SA-06b	Business Rule SA-06b: Data-in-transit shall be signed using NSA-approved signature algorithms (see [60]).	64
SA-06c	Business Rule SA-06c: Data that is transferred between security domains via physical transport of physical media (e.g., “sneaker-net”) shall be evaluated and sanitized IAW AR 25-2 [14] clause 4-16 and Information Assurance Best Business Practice Data Transfer Across Security Domains [60].	64
SA-06d	Business Rule SA-06d: Data shall not be electronically transmitted (“data-in-transit”) between security domains unless the CDS over which the data was transmitted was developed and approved IAW AR 25-2 [14] clause 4-21 and the Information Assurance Best Practice Cross Domain Solutions [62].	64
SA-06e	Business Rule SA-06e: The security markings of data transferred between security domains shall be updated for the new security domain as appropriate.	64
SA-07	Principle SA-07: Neglected, abandoned, or “residual” data is a security risk.	64
SA-07a	Business Rule SA-07a: Data that is moved, no longer useful, left on discarded devices, should be destroyed, wiped, purged, or sanitized IAW AR 25-2 [14] Section 4-20.	64
SA-08	Principle SA-08: The use of data service security guidance will make the incorporation of security controls into Army data services easier, improve the security of Army data, and reduce the costs of security solutions.	64
SA-08a	Business Rule SA-08a: Data stewards or an ADB designee shall develop, publish, promote, and maintain data service security guidance. The same body, or an allied body, should develop, test, and promote service security mechanisms, tools, and resources. The data service security guidance shall be derived from, complementary to, and consistent with Information Assurance access controls. The guidance shall include: <ul style="list-style-type: none"> • Security Information and Event Management (SIEM) policies/technology to improve data security by building-in the monitoring capabilities for threat prevention and auditing capabilities for intrusion forensics; 	65

Label	Principle/Business Rule	Page
SA-08b	<p>Business Rule SA-08b: Data service developers shall adopt and use data service security guidance in the design, development, fielding, and operation of data services. In particular, data service developers should adopt and use the following:</p> <ul style="list-style-type: none"> • WS-Security is a standard extension to SOAP to apply security to web services. The protocol specifies how integrity and confidentiality can be enforced on messages and allows the communication of various security token formats, such as SAML, Kerberos, and X.509. It uses XML Signature and XML Encryption to provide end-to-end security. • Security Assertion Markup Language (SAML) [63] is a standard for the exchange of principal authentication and attribute information between clients, services, and security services. • eXtensible Access Control Markup Language (XACML) [64] is a standard that defines a declarative access control policy language and processing model to evaluate authorization requests according to the rules defined in policies. • WS-SecurityPolicy [64] is a standard that can be used to enhance WSDL specifications to represent and exchange security policy information. • Tactical Service Security System (TS3) security handlers [66] is a software product developed by the Army for incorporation into web services development and deployment to support the consistent implementation of security mechanisms. TS3 incorporates WS-Security and SAML. • Identity and Access Management (IdAM) - Reference Architecture [67] specifies requirements for the implementation of Identity Management for authentication of principals and controlling access to data services. 	65
SA-09	<p>Principle SA-09: Authorization of principals to access data services via uniform access control mechanisms across the Army will ease availability of and access to data while maintaining a high level of security.</p>	65
SA-09a	<p>Business Rule SA-09a: Data stewards or an ADB designee shall establish, publish, and promote authoritative access control policies and mechanisms for data service access, including:</p> <ul style="list-style-type: none"> • Managing access by Attribute-Based Access Control (ABAC) policy to control the level of access by both anticipated and unanticipated users; • Developing and fielding an enterprise-level Policy Decision Service that is used by web services to authorize and authenticate access; and • Establishing an access control policy for each protected resource. 	65
SA-09b	<p>Business Rule SA-09b: Army security access control policies and mechanisms shall be used in the design, development, fielding, and operation of data services. In particular, data services shall comply with the access control requirements stipulated in AR 25-2 [14] Section 4-12.</p>	66
SA-09c	<p>Business Rule SA-09c: Security elements of DSRA [18] and the ADF-Data-related Security chapter [14] should be applied in planning and implementation of security measures.</p>	66
SA-09d	<p>Business Rule SA-09d: Data service design, development, and fielding shall follow NCES's security policies [60] [69].</p>	66
SA-09e	<p>Business Rule SA-09e: Data service design, development, and fielding shall follow the guidelines defined in the DISA STIGs [37].</p>	66

Label	Principle/Business Rule	Page
SA-09f	Business Rule SA-09f: Data services shall obtain a Certificate of Networthiness issued by NETCOM before fielding.	66
SA-10	Principle SA-10: Authorization to use a data service is distinct from authorization to access the data provided by the service. Security and protection of data is ensured by matching the authorization level of an authenticated identity with the security level of the data.	66
SA-10a	Business Rule SA-10a: A data service shall authorize data access by validating the security level (i.e., privileges) of the authenticated identity of a requesting consumer against the security level of the requested data.	66
SA-10b	Business Rule SA-10b: A data service should validate the “need-to-know” of the consumer requesting the data. Data service consumers and data service providers should adopt the XML Data Encoding Specification for Need-To-Know Metadata [70] for automating the validation of need-to-know.	66
SA-10c	Business rules SA-10c: A data service should exchange data IAW Intelligence Community Multi-Audience Tearline (IC-MAT) [71] when appropriate based on the content of the exchange.	66

Appendix D Relationship to Other Data Strategy Products

Table 5 identifies and describes Army and DoD data strategy products that are related to the AIA. Other products will be added to the list as they are identified.

Table 5: AIA Companion Products

Product	Description
DoD Information Enterprise Architecture [1] [2] [3]	The AIA structure and intent is based, in part, on that of the DoD IEA (particularly version 1.2 [1]). The DoD IEA is a DoD architectural guidance document that defines a common foundation to support DoD transformation to net-centric operations and establishes priorities to address barriers to its realization. The DoD Information Enterprise (IE) comprises the information, information resources, assets, and processes required to share information across the Department and with mission partners.
Common Operating Environment (COE) Architecture [4]	The AIA is a component of COE Architecture. The COE Architecture defines an approved set of computing technologies and standards that will enable secure and interoperable applications to be developed rapidly and executed across a variety of computing environments: data center/cloud, Command Post, mounted units, mobile devices, sensors and platforms, and real time environments. Each computing environment has a minimum standard configuration that supports the Army's ability to produce and deploy quickly high-quality applications, and to reduce the complexities of configuration, support and training associated with the computing environment.
Army Data Framework (ADF) [14]	The ADF combines components and patterns and provides guidance for developing a particular kind of information system. The ADF-Data Warehouse [19], for example, provides a standardized view of the components and organization of a Data Warehouse.
Data Strategy Reference Architecture (DSRA) [14]	DSRA is comprised of components and patterns that may be used in the design of data-centric systems; for example, an Enterprise Search pattern consists of a client who launches a search, a search engine or portal that performs the search, a search space that is the extent of the search, and potential delegated searches that launch other search engines over other search spaces. The DSRA provides the "ingredients" for the design of information systems, and the ADF provides the "recipes" for combining the ingredients to achieve particular goals.
Data Services Layer –Army (DSL-A) [18]	DSL-A is a framework and set of data service interface specifications that enables the Army and supporting organizations to develop data services that expose data assets, authoritative and otherwise, to consumers across the Army.
Namespace Management Enterprise Solution (NMES) [21] [22]	NMES provides an XML namespace management solution for the Army. It provides a reliable means to avoid XML namespace naming collisions on an Army scale. It provides a naming scheme for XML namespaces based upon a hierarchy of functions of interest to the Army.

Product	Description
<i>Rules for Cross-Cutting Capability (CCC) Information Exchange Specifications (IES) in Interface Specifications (“IES Data Rules”) [23]</i>	The IES Data Rules provide a drill-down focus on the use of IESs within the COE and across CEs to provide cross-cutting information sharing capabilities.

Appendix E Catalog of Data Standards

E.1 Data Standards

Data Standards address the data standard components of the Data Services and Infrastructure Services layers of the COE Implementation Plan [6] Technical Reference Model (TRM), illustrated in Figure 12. Complete definitions of the TRM layers are provided in the TRM itself. The Data Services and Infrastructure Services layers from the TRM layers are also illustrated in Figure 13 to show the relationship between the standards identified in this document and the COE Implementation Plan TRM.

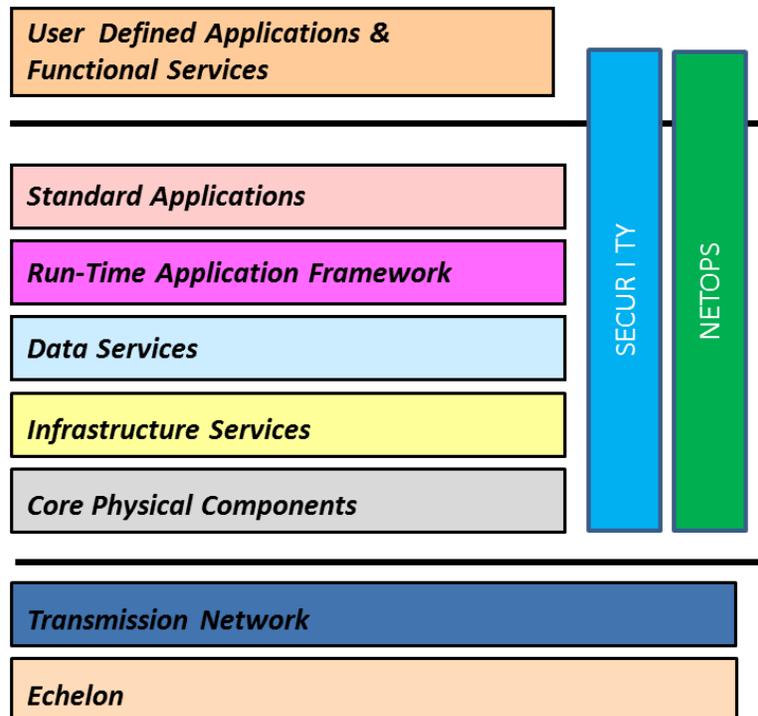


Figure 12: COE Implementation Plan Technical Reference Model (TRM)

Data standards fall into four broad categories as illustrated in Figure 13. The most fundamental standards are those foundation data standards dealing with bit/byte level patterns for representing primitive information and structure. These standards are widely known and used through the information technology industry; some of these will be listed in this document, though many of the most general standards (e.g., ASCII) are not.

The second category of data standards is infrastructure data standards that pertain to technology and its use. These standards are applicable throughout the COE and are not tied to or unique to any specific domain.

The third category of data standards consists of those that specify the exchange format for usage-domain-specific information – they are “about” something other than information technology.

The fourth category of data standards are data access standards. These standards do not deal with information content or data structuring requirements, but focus on the access methods for retrieving data.

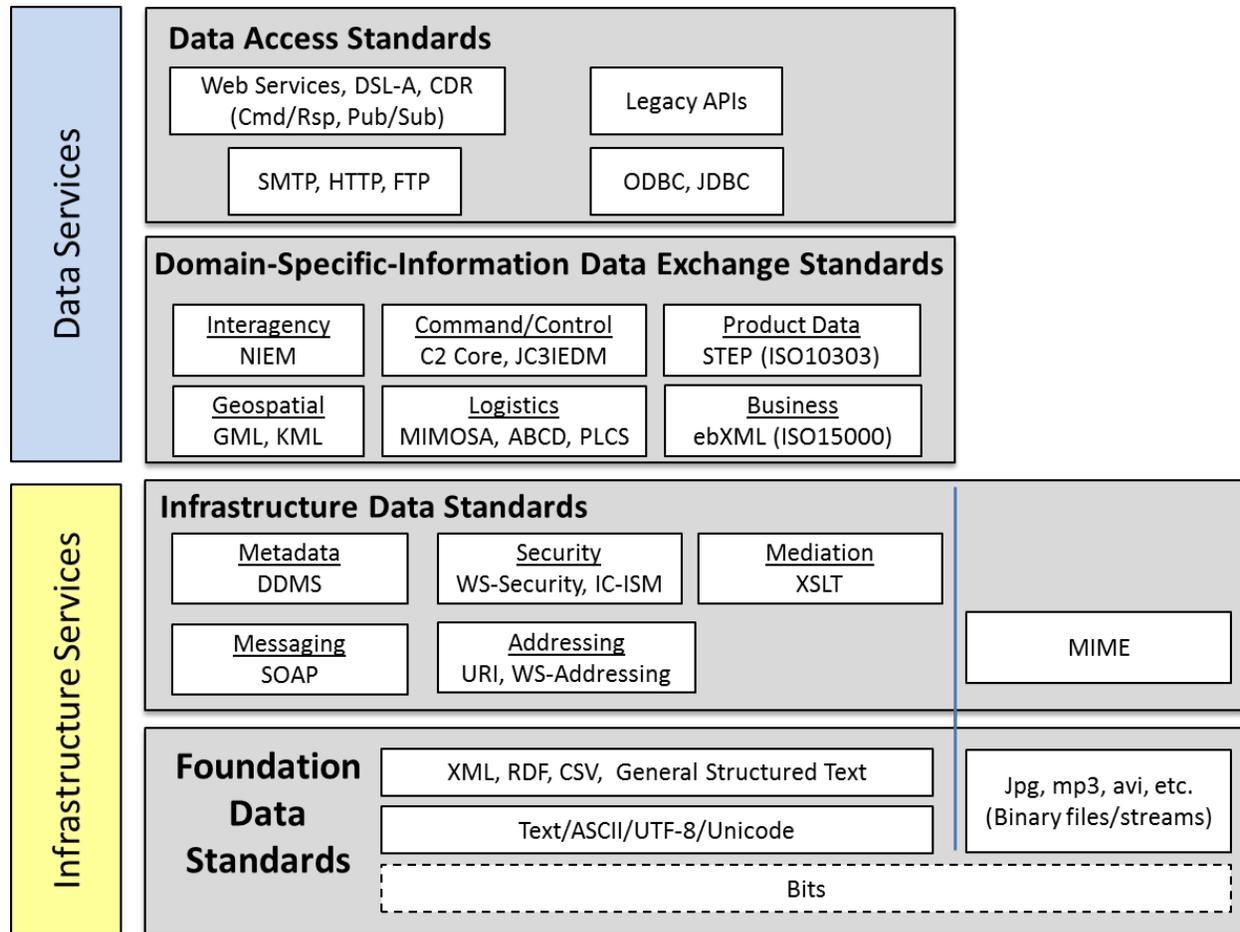


Figure 13: Data Standard Classifications

In the tables listed below, the following legend shall be used to denote the computing environments in which the standards are applicable:

Computing Environment Key:

- DC = Data Center
- CP = Command Post
- MN= Mounted
- MB = Mobile
- SN = Sensor
- RT = Real Time

The DISR Status column denotes whether the standard is:

- M = Mandatory
- E = Emerging
- N/A = Not in DISR

E.2 Foundation Data Standards

Table 6: Foundation Data Standards

Standard ID/Title	Description	DISR Status	Computing Environment					
			DC	CP	MN	MB	SN	RT
XML Extensible Markup Language	Basic platform-independent data formatting and structuring	M	X	X	X			
ODBC/JDBC	Database connections internal to a system for data access and storage. Not intended to exposure to external clients.	N/A						

E.3 Infrastructure Data Standards

Table 7: Infrastructure Data Standards

Standard ID/Title	Description	DISR Status	Computing Environment					
			DC	CP	MN	MB	SN	RT
WC3 URI; WS-Addressing	Web Addressing	E	X	X	X			
WC3 SOAP, Part 1 and 2	Service Messaging	M	X	X	X			
WS-Eventing and WS-Notification	Service Interaction (pub/sub)	E	X	X	X	X		
W3C RESTful; WS-Transfer	Service Interaction (Create Read Update Delete)	N/A	X	X	X			
WS-Security	Secure Services	M	X	X	X	X	X	
XSLT	XML data transformation/mediation	M	X	X				
UDDI	Service Discovery	M	X	X				
DDMS	DoD Discovery Metadata Specification	M	X	X				
XML Schema Part 1 and 2	Data Model Specification for XML documents	M	X	X				
DOM Level 3; XQuery; XPath	XML document processing	M	X	X	X			
IC-ISM.XML	Dissemination Controls	M	X	X				
Transport-Level: TLS Message: XML Signature	Integrity	M	X	X	X	X		
Transport-Level: TLS Message: XML Encryption	Confidentiality	M	X	X	X	X		
WS-Security Attribute Services ABAC	Authorization/Access Control	N/A	X	X	X	X		

Standard ID/Title	Description	DISR Status	Computing Environment					
			DC	CP	MN	MB	SN	RT
Host-Based Security System (HBSS)	Security for Data at Rest	N/A	X	X				

E.4 Domain-Specific-Information Data Exchange Standards

Table 8: Data Exchange Standards

Standard ID/Title	Description	DISR Status	Computing Environment					
			DC	CP	MN	MB	SN	RT
Command and Control / Tactical Standards								
C2 Core Command and Control Core (C2 Core)	The C2 Core vocabulary explicitly defines C2 concepts which are commonly used across the entire Joint C2 domain, from the global/strategic to the tactical levels of operation, and specifies the format/structure for exchange C2 data	N/A	X	X	X	X		
JC3IEDM Joint Consultation Command and Control Information Exchange Data Model	Enables international interoperability of C2 information systems at all levels from corps to battalion (or lowest appropriate level) in order to support multinational (including North Atlantic Treaty Organization (NATO)), combined and joint operations. JC3IEDM focus on information that supports: Situational awareness; Operational planning; Execution; Reporting.	N/A	X	X	X	X		
GFIEDM Global Force Management Information Exchange Data Model	Global Force Management Person Type Skill Attributes	N/A	X	X	X	X		
MIL-STD-6040B(C1) United States Message Text Format (USMTF)	Provides U.S. goal-allied adoption of MTF standards that enables interoperability among NATO and other U.S. and allied C2 Information Exchange Systems.	M	X	X	X	X		
MIL-STD-6017B VMF Variable Message Format	VMF (MIL-STD 6017) messages provide a common means of exchanging digital data across any interface between combat units at various organizational levels with varying needs for volume and detail of information and are applicable to a broad range of tactical communications systems.	M			X	X	X	X
BML Battle Management Language	BML specifies the data that conveys information used to command and control forces and equipment conducting military operations. Provides for situational awareness and a shared, common operational picture. It is intended as a representation of a commander's intent to be used for real troops, for simulated troops, and for future robotic forces	N/A	X	X	X	X		

Standard ID/Title	Description	DISR Status	Computing Environment						
			DC	CP	MN	MB	SN	RT	
Interagency Standards									
NIEM National Information Exchange Model	The National Information Exchange Model (NIEM) is a data exchange standard intended to facilitate the exchange of information among federal, state, local, and tribal agencies, as well as with private sector entities. NIEM covers a number of diverse subject areas: Chemical, Biological, Radiological, Nuclear (CBRN); Emergency Management; Immigration; Infrastructure Protection; Intelligence; International Trade; Justice; Maritime; Screening; (Youth and) Family Services.	N/A	X	X					
Geospatial Standards									
GML Geography Markup Language	Defined by the Open Geospatial Consortium (OGC) to express geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet.	M	X	X	X	X	X	X	
OGC KML	Expressed geographic annotation and visualization on Internet-based, two-dimensional maps and three-dimensional Earth browsers.	M	X	X	X	X			
Product Data Standards									
ISO 10303 (STEP) Standard for the Exchange of Product model data	Representation and exchange of product information for manufacturing, acquisition, logistics, or other domains in which physical product information must be exchanged.	M	X	X					
Logistics Standards									
ISO 10303-239 (PLCS) Product Life-Cycle Support	A volume of ISO 10303. Representation and exchange of information required for logistics support of high-value, long-life systems (e.g., aircraft, naval vessels).	N/A	X	X	X	X			
OSA-CBM (MIMOSA) Open System Architecture for Condition-Based Maintenance	Specifies the information required for interoperability of Condition-Based Maintenance systems and the mechanisms for moving the information between systems.	M	X	X	X			?	
ABCD Army Bulk CBM+ Data	The ABCD file specification integrates meta-data about the platform and sensor locations, the missions/regimes performed, and export file information utilizing specific Machinery Information Management Open Systems Alliance (MIMOSA) Common Relational Information Schema (CRIS)-defined metadata into the Common Data Format (CDF) scalar and/or multidimensional data.	N/A	X	X					
CRIS Common Relational Information Schema	Logistics actionable data standard	N/A	X	X	X			X	

Standard ID/Title	Description	DISR Status	Computing Environment						
			DC	CP	MN	MB	SN	RT	
Business									
ISO 15000 (ebXML) Electronic Business using eXtensible Markup Language	A family of XML-based standards that provides an infrastructure for electronic (digital) business information interoperability.	N/A	X						

E.5 Data Access Standards

Table 9: Data Access Standards

Standard ID/Title	Description	DISR Status	Computing Environment						
			DC	CP	MN	MB	SN	RT	
Content Discovery & Retrieval (CDR) Standards [20]									
CDR REST/SOAP Search	Specifies standard service interface for searching through content and metadata in both structured and unstructured search domains. There are specifications for both SOAP-based and REST-based interfaces.	M	X	X					
CDR REST/SOAP Brokered Search	Specifies standard service interfaces for conducting a search across multiple search services and returning an aggregate response. There are specifications for both SOAP-based and REST-based interfaces.	M							
CDR REST/SOAP Retrieve	Specifies standard service interface for retrieving an identified content resource and deliver it to requestor. There are specifications for both SOAP-based and REST-based interfaces.	M							
CDR REST/SOAP Deliver	Specifies standard service interface for delivering a content resource to specified location. There are specifications for both SOAP-based and REST-based interfaces.	E							
Data Services Layer - Army (DSL-A) Standards ⁷ [18]									
DSL-A Retrieve Pattern Interface Specification	Specifies standard service interface for retrieving data. It is agnostic with respect to the structure of the data.	N/A	X	X					
DSL-A Modify Pattern Interface Specification	Specifies standard service interface for creating, updating, and deleting data. The Retrieve and Modify patterns together specify common database CRUD functionality.	N/A	X	X					
DSL-A Search Pattern Interface Specification	Specifies standard service interface for searching for data.	N/A	X	X					

⁷ The DSL standards listed are key DSL services; the list is representative and not exhaustive.

Standard ID/Title	Description	DISR Status	Computing Environment					
			DC	CP	MN	MB	SN	RT
DSL-A Data Access Service Interface Specification	Combines Retrieve and Modify pattern capabilities.	N/A	X	X				
DSL-A Federated Search Service Interface Specification	Provides search capabilities across a federated collection of search spaces.	N/A	X	X				
DSL-A Transform Pattern Specification	Specifies standard service interface for transforming data. The specification makes allowances for other transformation specifications but focuses primarily on XLST transformation.	N/A	X	X				

Appendix F Data Services

F.1 Data Service Family Descriptions

Data services are the subset of the services that comprise a SOA system design that collectively provide “Data as a Service” to application level services and workflows. They are “lower level” services in that they are closer to actual data assets. In addition to basic Search-Create-Read-Update-Delete (SCRUD) capabilities, they provide user-oriented data functions such as query management and data dissemination as well as “back office” functions such as archival, replication and change control.

For convenience and understanding, data services are grouped into service categories and service families. Service categories correspond to the “user-level” and “back office” groupings:

- *Enterprise and Provider Data Services* are services directly used by consumers and consuming applications to provide end-user functionality.
- *Data Management Services* are the “back office” services that provide data management capabilities that are not typically seen or used by end-users.

Data services are further categorized by function in Table 10 and Table 11.

Table 10: Consumer-Accessible Data Services

Consumer-Accessible Data Services	Description
Discover	A service that enables a consumer to search for data and/or data related services.
Access	A family of services that enable authorized users or applications to retrieve, and modify data in accordance with the permissions assigned to their role or activity.
Mediate	A family of services that enables a consumer to use data from other services and to produce a coherent, usable set of information, making use of translation, transformations, or simple semantic mappings and validation. The use of a neutral mediating form can reduce the n-squared problem.
Disseminate	A service that moves data to one or more designated consumers.

Table 11: Data Management Services

Data Management Services	Description
Ingest	A service that automates loading bulk data into an application or system. In limited cases, it may also include some initial, albeit minimal, pre-processing of the data received.
Archive	A service that supports saving data to long term secure storage to support historical or other uses. This generalized category includes backup and restore.
Replicate	A service that provides the capability to ensure consistency between local copies of persisted data distributed across a network in order to improve reliability, fault tolerance, or accessibility.
Store	A service that provides the capability to store data in a persistent manner (in non-volatile memory or other data storage media).

F.2 Data Service Descriptions

The functionality of the individual data services in the CDR [20] specifications are described in Table 12.

Table 12: CDR Data Service Descriptions

Service Category	Service Name	Description
Provider Data Services		
Access	Retrieve	The Retrieve Service enable consumers to retrieve of an identified content resource from a Content Collection in which it is stored and initiate delivery of the retrieved resource to the requestor or to a designated alternate location using the Deliver Service.
Access	Deliver	The Deliver Service complements the Retrieve Service enabling a content resource to be delivered to a specified location which may or may not be the requesting consumer. It provides additional processing of the content to make it suitable for delivery to its destination and delivery path to be used. On behalf of the requesting consumer, the service may also retrieve the requested content and then deliver to the specified location.
Access	Query Management	The Query Management Service enables a consumer to create, update, and store the search requests as Saved Searches to execute Saved Searches based on their specific request or on event triggers.
Discovery	Search	The Search Service enables consumers to search through content and metadata in multiple formats as specified by the consumer, such as image files and textual documents. It also enables searching through information content that is static, dynamic, structured and unstructured; and searching through and appropriately processing of information content and metadata at different classification levels, and with different handling caveats; information which could be located on different security domains. It also enables searching through natural language content (probably in many different languages) or highly formatted content such as geospatial or temporal content.
Discovery	Brokered Search	The Brokered Search Service facilitates the distribution of queries to applicable/relevant Search Services and content collections these Search Services expose. It aggregates the results returned individually into a single, uniform results set which is returned to the requesting consumer.
Discovery	Describe	The Describe Service is a complement to Search that enables resource providers to expose information describing their content collections and content resources. It provides interested parties with a description of the resource and how it can be accessed or used.

The functionality of the individual data services in DSL-A [18] is described in Table 13.

Table 13: DSL-A Data Service Descriptions

Service Category	Service Name	Description
Provider Data Services		
Access	Retrieve	The Retrieve Service provides consumers the ability to retrieve data or artifacts from data assets on the network.
Access	Modify	The Modify Services provides consumers with appropriate credentials to create, update, or delete data from data assets.
Access	Deliver	The Deliver Service complements the Retrieve Service by providing a consumer the ability to re-route the requested data or artifact to a destination other than the point of request. This service would be used, for example, by a consumer that wants to request the retrieval of a very large file from a mobile device but receive that file at a desktop computer.
Access	Query Management	The Query Management Service enables a consumer to create, manage, and share complex data retrieval requests, i.e., "canned queries." This service is important in the development of, for example, dashboards and reports.
Discovery	Search	The Search Service provides consumers the ability to search for data or artifacts across the network in the same way that people use commercial search engines to search the internet. The Retrieve and Search Services together provide the most basic and widely-used functionality by data consumers.
Discovery	Brokered Search	The Brokered Search Services enables consumers to conduct a search across a set of separate search spaces and conduct the search asynchronously (i.e., consumer submits a search and then "walks away" and awaits delivery of the search results at some later point in time.)
Discovery	Data Service Discovery	The Data Service Discovery Service enables consumers to search for other services that provide a particular kind of data or a particular kind of functionality. This service would typically be used by consumers developing an application or creating a workflow.
Discovery	Registration / Describe	The Describe Service is a complement to Search that enables a data provider to describe a data asset (i.e., create metadata and make it available) such that a Search Service can consume the description and, thereby, make searches for the data asset faster and more accurate. The Registration Service complements Describe by providing the capability to explicitly register and describe the data asset in a public registry.
Mediation	Data Mediation	The Data Mediation Service provides consumer services the ability to transform data from one governing format to another, e.g., via an XLST script.

Service Category	Service Name	Description
Dissemination	Data Dissemination	The Data Dissemination Service enables data provider to distribute or broadcast data to a consuming audience.
Dissemination	Publish/Subscribe	The Publish/Subscribe Services provide complementary consumer/provider capabilities that enable a consumer to “express an interest in” a data item, artifact, or event by subscribing to it, and a provider to broadcast changes to or information about the data/artifact/event.
Enterprise Data Services		
Storage	Transaction	The Transaction Service, like common database transaction management, provides the Modify Service the capability to successfully complete a complex or time-consuming set of CRUD actions or roll-back the actions if they cannot be completed.
Storage	Change Control	The Change Control Service complements the Modify service by tracking changes to data in a data asset by monitoring create, update, and delete requests.
Storage	Data Storage	The Data Storage Service provides the capability to persistently store data at a reliably accessible location on the network.
Ingest	Data Ingest	The Data Ingest Service provides a data consumption (and possibly pre-processing) capability that is typically used to consume and load bulk data into, for example, a data warehouse.
Archive	Data Archive	The Data Archive Service complements the Data Storage Service by provide the capability to move data to safe and secure archival storage when no longer needed on a regular basis.
Replication	Data Replication	The Data Replication Service provides the capability to duplicate or mirror a data asset at a separate location and keep the data assets synchronized. The capability is typically needed when the latency involved in data access from a particular consuming location is unacceptably large or problematic (e.g., Afghanistan accessing CONUS-based data); the latency is improved by physically moving the data closer to the point of access.

The DSL-A Service Interface Specifications can be found at:

<https://www.intelink.gov/inteldocs/view.php?fDocumentId=342142>

Appendix G Processes and Activity Models

G.1 Community of Interest Processes

A COI is a group of users with a set of shared goals, interests, missions, or business processes and a need to collaborate in pursuit of the shared goals, etc. This group includes end users, program managers, application developers, subject matter experts, Combatant Command, Service and Agency representatives, and IT Portfolio representatives. COIs are an approach for developing the agreements necessary for the development of interoperability solution and information sharing. The COI concept is described in the DoD Net-Centric Data Strategy [31] and directed by DoD Directive 8320.02 [7].⁸

There are numerous COIs that are registered at the DoD DSE 2.0 COI List web page that have a status that ranges from Proposed to Effective. Any organization can join these COIs as long as they have shared goals, interests, missions or business processes.

The creation of a new COI is only advised when there are no other COIs with shared goals, interests, missions or business processes. A COI having a clearly stated, well defined purpose and scope tends to be successful in accomplishing its goals. The recommendation is for COIs to form with a 'top-down' authority and an official charter. Advantages of forming a COI using the "top-down" authority approach are having a clearer vision of the scope and goals, and creating a shared vocabulary that accommodates all the participants' information needs up front resulting in a reduction of complications when sharing information.

Advantages of Forming a New COI versus Joining an Existing COI

- The COI's goals will be well aligned with the organization's goals.
- Priorities and schedules will be better aligned with the organization's requirements.

Disadvantage of Forming a New COI versus Joining an Existing COI

- There is significant time and cost overhead in setting up a COI and its governance process. A considerable amount of the initial time would be spent executing the steps listed below in. If the organization joined an existing COI, these steps might have already been performed, and collaboration could begin almost immediately.

Procedure

The recommended procedure for forming a new COI is as follows:

1. **Define the COI scope.** Successful COIs are able to define and maintain a tight scope and focus. A COI scope, preferably one sentence, should describe the information sharing problem that the COI will address.
2. **Advertise the COI.** This will ensure that DoD users can discover its existence and mission, and give them the opportunity to participate. Register the COI in the DoD COI directory by clicking the "Add a COI" link on the bottom right of the following web page:
3. **Identify the COI membership.** Initial membership will coalesce around a common mission and information sharing problem. Members are those who stand to benefit and those whose processes and/or systems will change as a result of COI activities.

⁸ Adapted from <http://cio-nii.defense.gov/sites/coi/governance.shtml.htm>

4. **Establish COI governance.** Establish a charter. The COI governance structure establishes a decision-enabling framework that directs and controls the COI and assigns accountability to support the mission. A charter should be used if there could be an issue about the allocation of resources, such as the contribution to the COI from several PEOs or several PMs.
5. **Kickoff COI.** The kickoff meeting will discuss the previous steps and is accompanied by the COI Plan of Action and Milestones (POA&M). A sample COI kickoff slide deck can be found in the documentation referenced below.

Additional information about governance and the kickoff for a COI is available at:

COI Governance and Guidance:

<http://dodcio.defense.gov/CommunitiesofInterest/COIGovernanceandGuidance.aspx>

COI Kickoff Template:

<http://cio-nii.defense.gov/sites/coi/Training/DT-05-Sample-COI-Kickoff-Template.ppt>

COI Kickoff Example

<https://www.intelink.gov/inteldocs/view.php?fDocumentId=178437>

G.2 Data Planning Processes

G.2.1 Authoritative Data Source Processes

The process for establishing an ADS is presented in the *DSE Concept of Operations* [60] and is illustrated in Figure 14 (from [60]).

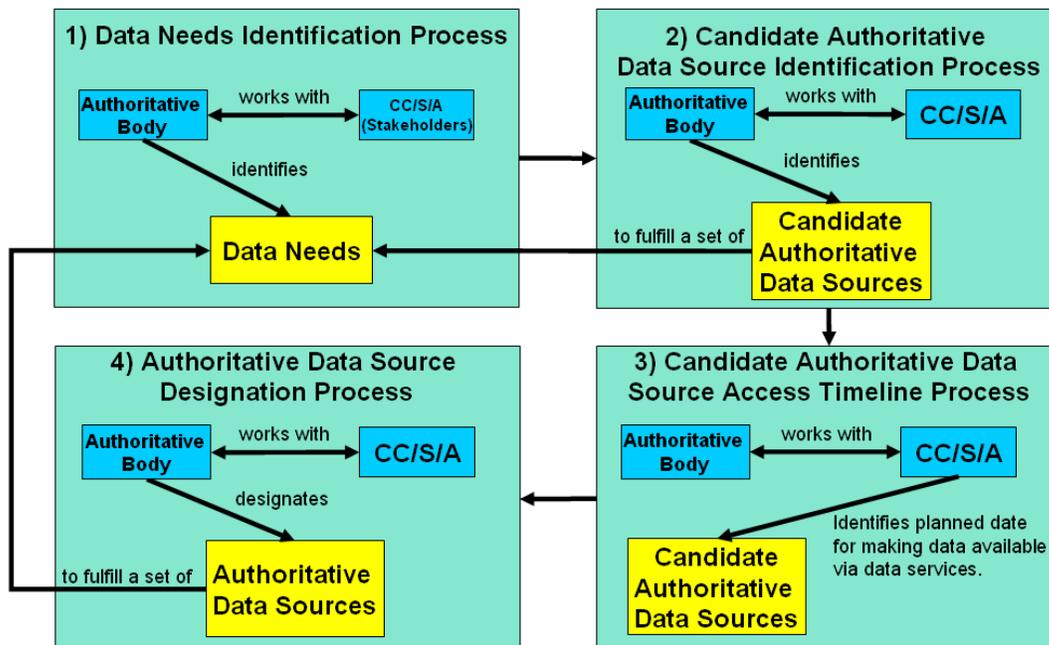


Figure 14: Process of Establishing an Authoritative Data Source in DSE 2.0⁹

⁹ CC/S/A: COCOM/Service/Agency

The Army ADS registration process that compliments the DSE 2.0 ADS process is described at https://www.milsuite.mil/wiki/Authoritative_Data_Sources_Process

G.2.2 Information Exchange Specification Processes

The development and maintenance of the IES follows the DSRA Information Architecture [18] Data Reference Model (DRM) element called Information Exchange Standards Specification (IESS). This section summarizes the DSRA process.

Creating the IES follows the DSRA pattern *Creating Data Exchange Specifications* and consists of three primary steps:

- **Specification Development:** Developing a normative technical specification for referencing (requiring conformity to) and for conformity (claims of conformity to).
- **Harmonization:** Reducing incompatibilities with existing systems and specifications.
- **Optimizing Reuse/Modularity:** Making best use of existing specifications (i.e., not rewriting them) and allowing the technical specification to be used by the community (affording other specifications layered on top).

The contents/composition of an IES is described in Section 7.2.2.

Developing an IES requires the cooperation and participation of community members that will be using the IES. The cooperation is described by the DSRA pattern *Agreements Among Organizations*. This pattern is comprised of the following activities:

- **Establishing a Contract:** The agreement becomes a contract among the trading partners.
- **Agreed upon Meaning:** The parties agree upon the meaning of the data exchanged. **Note:** for example, parties might agree to use an "invoice standard" as part of their data exchange, however one set of parties use the invoice standard for "invoice printing" (i.e., party X is an invoice printing service where party Y asks X to print Y's invoices), whereas another set of parties use the invoice standard for "commercial invoicing" (i.e., party P tells party Q that certain monies are owed for products/services).
- **Referencing a Technical Specification:** The agreement references a data exchange specification that describes the data transferred and the meaning of that data (within the scope of the data exchange).

The maintenance of the IES follows the DSRA pattern *Lifecycle Management of Technical Specifications*. The lifecycle is illustration in Figure 15.

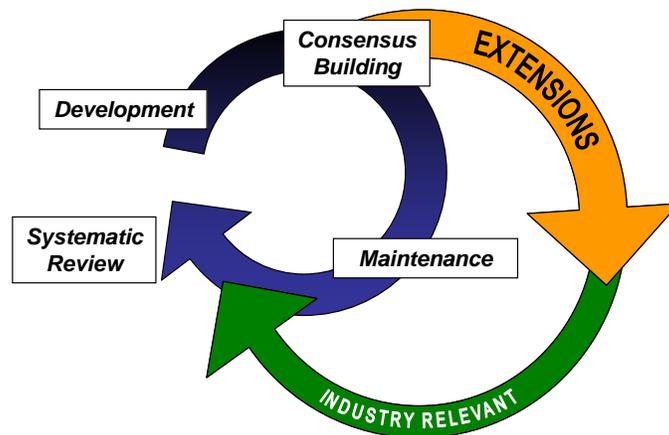


Figure 15: Lifecycle of Technical Specifications

The lifecycle process for the development and maintenance of technical specifications consists of the following steps:

- Specifications (and standards) are developed.
- Specifications are approved via a consensus-building process.
- Once approved, the specification is maintained, i.e., process defects (technical corrigenda), formal interpretation, and amendments.
- After approval, users/industry develops/experiments with “extensions” to the specification (e.g., new technology, new features).
- Some of those “extensions” (industry/COI relevant ones) might be considered in the next revision of the specification.
- At some point after approval (typically, 3-5 years), specifications are reviewed with three possible outcomes: reaffirm (no changes, specification is still relevant), revise (specification to be improved), withdraw (no longer relevant technology).

G.2.3 Unstructured Data Processes

Unstructured data provides key intelligence that is needed to support the Army’s Intelligence, Surveillance and Reconnaissance (ISR) tasks. Unstructured data needs to be processed to make it Visible, Accessible, Understandable and Secure (VAUS) to potential users. The key is to provide structure to unstructured data. Several Army organizations are addressing this issue. An example is the DCGS-A Secure Internet Protocol Router (SIPR) Cloud System (DSC) which is a large scale data storage, processing, and integration (DSPI) system that works regardless of the characteristics of the data (modality, structure, or representation) and rigorously enforces information security and privacy controls so that the entire Intelligence Community (IC) can cultivate and exploit all data, information and knowledge. The DSC implements a solution that meets the data sharing Intelligence Community Directive (ICD) 503 [74], particularly, the ability to:

- Search across source data boundaries
- Enrich across source boundaries (asset new elements and associations)
- Reuse/repurpose data
- Not lose / distort data or semantics

The DSC Dataspace data architecture is divided into three segments (or logical layers) and is targeted to support both structure and unstructured data:

- Segment 1 (ADF=Artifact Data Framework) serves as a unified interface to the sources; it stores the sources metadata and in some cases can store the content of the unstructured source.
- Segment 2 (DDF=Data Description Framework) serves as a unified structured store for data and data-semantics of the structured sources, and disambiguated data from the unstructured sources. The storage model is defined by the DDF¹⁰.
- Segment 3 (MDF=Model Description Framework) serves a unified store for the data-models of the sources. In the current document this layer is discussed only to the extent needed to maintain and manipulate Segment 2.

The DDF is an abstraction over (data) models; thus DDF can capture any model and be implemented in any meta-model. The DDF divides the data architecture into a series of concepts:

- **Source:** The origination system/application for artifacts (i.e. data). Humans can also be sources of artifacts.
- **Artifact:** A tangible container of source data; many formats are supported, including:
 - Data in a file
 - Data in a relational database
 - Data retrieved via an interface from legacy application/system
 - Real-time data feed
- **Mention:** A chunk of data located within a tangible artifact at a quantifiable span (location).
- **Sign:** A representation of all mentions those are identical except for their indexicality (across source and location).
- **Concept:** A representation of an abstract idea, defined explicitly or implicitly by a source data model. For example, the nodes of an ontology, the tag set in an XML Schema Document (XSD), and the attribute /table names in a relational database all represent concepts.
- **Term:** A disambiguated sign abstracted from the source artifact or asserting analyst. The process of disambiguation associates a sign with a concept using the `isInstanceOf` predicate.
- **Predicate:** A representation of an abstract idea, i.e. a concept, used to express a relationship between “things.” Predicates are used in the formation of statements (described below) and may be defined either explicitly or implicitly by a source data-model.
- **Statement:** Encodes a binary relationship between a subject (term) and an object mediated by a predicate. A statement is represented by an ordered triple {subject, predicate, object}.
- **Metadata:** There are various types of metadata:
 - **Artifact Metadata:** Metadata that describes the properties of the particular instance of an artifact/data (e.g., originating system/application, creation date, ingestion date, size, Multipurpose Internet Mail Extensions (MIME) type, language, character set, etc.)

¹⁰ <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4753515>

- **Source Metadata:** Metadata that describes the contents of a data asset. For example, in the case of an email source, the “from”/”to” tags are metadata; in the case of a relationship source, it is the data model.
- **Term and Statement Metadata:** Metadata describing the term/statement (e.g., author, date).
- **Concept and Predicate Metadata:** Metadata describing the concept/predicate (e.g., author, date).
- **Text/Geo Index:** The mentions of an artifact can be directly searchable using a keyword or geo coordinate.

These DDF concepts are then represented as a series of tables and relationships in the DSC data model. These are then available via the DSC interfaces. Figure 16 illustrates the relationships among the Data Description Language (DDL) concepts.

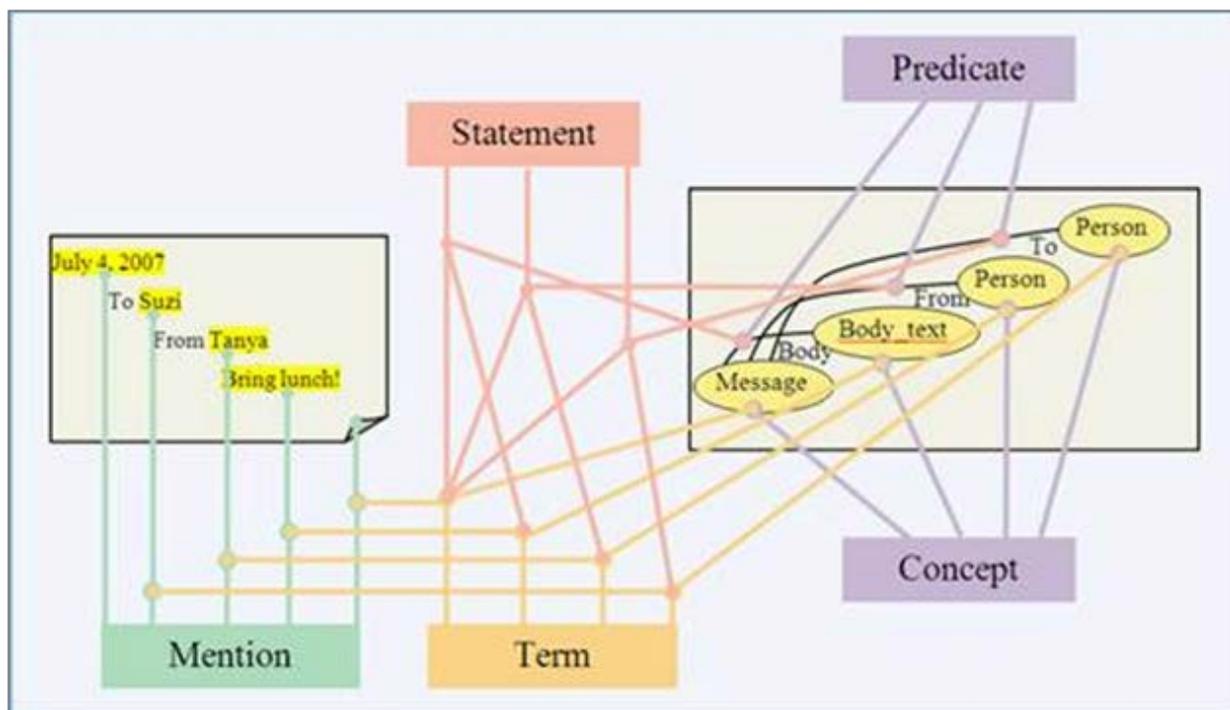


Figure 16: Rainmaker DDF (Conceptual)

Implementing a solution like the DSC Dataspace data architecture is needed to expose both structured and unstructured data. Organization will need to establish policies to ensure that data is managed and processed to make it VAUS to potential users.

DCGS-A Cloud and Data Strategy Resources:

DCGS-A Standard Cloud Intelink WikiPage

https://www.intelink.gov/wiki/DCGS-A_Standard_Cloud

Army G2 CIO Interview on “DoD deploys high-tech tools to boost intel collaboration”

http://defensesystems.com/Articles/2011/05/23/QA-Lynn-Schnurr-Army-intelligence.aspx?s=ds_230511&admgarea=TC_DEFENSE&Page=1

G.2.4 Data Model Planning and Organization

G.2.4.1 Function, Role, and Purpose of Data Models

The term “data model” has many senses of meaning and is often prepended with an adjective to denote one of these senses, such as:

- *Conceptual* data model, meaning a data model that represents the general concepts within a domain and ignores many of the details needed to describe those concepts;
- *Logical* data model, meaning a data model that represents an abstract design view of data that has not been tailored (e.g., denormalized) to make it more suitable for a particular software application; and
- *Physical* data model, meaning a data model that directly and explicitly governs the format and structure of data (and is synonymous with the term “schema”).

This section will focus on the role of physical data models in information sharing and data exchange. Conceptual and logical data models may play a role in the development of physical data models, but addressing this process and relationships is beyond the scope of this section.

When used in its general form in this section, “data model” may refer to any of the above mentioned kinds of data models.

Physical data models are used to fulfill a number of different roles and perform different functions in COE implementations. Within a CE or a mission area involving many CEs, there will be different physical data models fulfilling these roles/functions. This section provides guidance on planning and organizing the physical data models within a CE or mission area.

Some functions are common to all physical data models; a physical data model:

- governs the structure of one or more datasets;
- provides a guide to the meaning of data in a dataset;
- has a business-functional purpose; and
- is “about” some real-world domain¹¹ (which establishes the scope of the data model).

Understanding what a data model is “about” (i.e., its scope) is critical in the planning and organization of data models (and, ultimately, interoperability) because data models are often “about” the same real world things. Data governed by a physical data model represents information about the real world domain. It is important to recognize that computer programs, files, and data are also real world things, and there may be data models “about” them (e.g., metadata).

The business-functional purpose of a physical data model is what the data model is “supposed to do.” For example, the business-functional purpose of IC ISM is to annotate data-in-transit with security information to facilitate data security (and IC ISM is “about” data security levels and data-in-transit.) The business-functional purpose of a physical data model may be simply to represent and exchange information about a real world domain, e.g., GML/KML represents geospatial information.

Within a CE or mission area, physical data models may be fielded in the following roles:

- governing data-in-transit; and
- governing data-at-rest.

¹¹ A “real world domain” is defined by the intersection of some or all of the following: time period, spatial/physical location or region; things, relationships between things, or processes in which things participate.

Data-in-transit is a message or stream of data moving over a network and is transient. Data-at-rest is data that is stored in databases or file systems and is persistent (stored messages are data-at-rest because they are persistent). The term “common data model” is often used to refer to physical data models governing data-in-transit between members of a known community of systems; it is also sometimes used to refer to a data model that provides an integrated or composite view of a collection of datasets at rest (which is also sometimes called a “logical” data model).

The role/function of data-in-transit is different than data-at-rest. Data-at-rest represents a persistent body of information about a real world domain; the body of information is a resource for applications or consumers. Data-in-transit is information sharing (i.e., an act of communication) and represents only a subset of information about a real world domain that is pertinent to the purpose of the information sharing act. The same data model may govern data-at-rest and data-in-transit, but the information represented by the latter is a subset (typically a much smaller subset) of the information presented by the former.

The design requirements for physical data models governing data-in-transit and data-at-rest are significantly different. Narrow-bandwidth, disadvantaged networks will require small, compact datasets and simple data models that are narrowly focused on scope and business purpose; the characteristics of inter-CE control points will affect the design requirements for data-in-transit data models. Data-at-rest in a data center will require larger, more complex physical data models with broader scopes and wider, more numerous business purposes.

Physical data models are also used to specify the message content of data service interfaces (e.g., the XML Schemas associated with a WSDL), fulfilling a data-in-transit role.

G.2.4.2 Planning and Organizing Data Models in a CE or Mission Area

Establishing the roles, purposes, scope, and functions of physical data models as described above is essential to planning and organizing the data models used with a CE/mission area. A single Information Exchange Specification governing all the data within a CE/mission area is an idealistic goal that is not, unfortunately, realistic or practical. In reality, there will be many different models that vary along the role/purpose/scope/function axes described above. Identifying and characterizing the physical data models along these axes is a necessary first step to fielding data models in COE implementations in CEs/mission areas.

Data models that govern data-at-rest are internal physical data models and are outside the scope of COE implementation guidance. However, the physical data model associated with a data service (i.e., the data service schema) can serve as a surrogate for the data-at-rest in the data asset “behind” the data service for the planning and organization purposes. (In Federated Database terminology, the data service schema is called an “export schema.”) It is a surrogate for the data-at-rest because it is “about” the same domain as the data asset and the information available through the interface is (presumably) the same information contained in the data asset. When the service is called, however, the data service schema behaves like a data-in-transit data model because the service request and response are data-in-transit.

Planning and organizing the data models in a CE/mission area involve the following steps:

- (1) Identify the data service schemas for the data services providing access to data assets within the scope of the CE/mission area. This may be a long list.
- (2) Ensure that the data service schemas in the list are properly annotated with metadata and registered at appropriate registries (e.g., the DoD DSE 2.0).

- (3) Document the scope (what the data model is “about”) and purpose (what the data model is “supposed to do”) of each data service schema. This step is essential for understanding the information available within the CE/mission area, identifying and planning ADSs, and harmonizing data models.
- (4) Identify high-value/high-frequency information exchanges within/across the CE/mission area. This step is necessary to design or select data-in-transit data models (i.e., physical data exchange schemas or Information Exchange Specifications) for transmitting information between points in the CE/mission area. This step requires an understanding of the business processes or missions that the fielded CE(s) must support.
- (5) Assess/determine the need for common data-in-transit data models. Assuming as a starting point that all data service schemas are different, this activity considers and answers the following questions:
 - a. Can consumers tolerate and effectively use multiple data service schemas? If so, then data service schemas are also data-in-transit data models.
 - b. Can/should the data service schemas be harmonized in accordance with an Information Exchange Specification or data model fragments? This would provide consumers more commonality across data service schemas, but will still result (from a practical perspective) in multiple data-in-transit data models that may share common fragments.
 - i. Mandating the use of a common data service schema would not be practical because the data model would effectively have to represent the union of all the information available from all the data assets – a common enterprise data model. This kind of model is impractical to develop and use.
 - c. Can one (or more) common data-in-transit data model be identified or developed to provide consumers with a single consistent view of information? Can data service schemas be mapped to (and data translated/mediated into) this Information Exchange Specification format?
 - d. Do CE requirements (e.g., disadvantaged networks) impact the design of the data-in-transit data models? How is data translated/mediated into this format?
- (6) Ensure that all information that must be exchanged within the CE/mission area is represented by at least one data-in-transit data model.
- (7) Identify information gaps in the available data service schemas. Information gaps are information that is needed by some agent within the CE/mission area and is not provided or available through a data service schema. This step should include the planning and design action necessary to fill the gaps.
- (8) Where translation/mediation is necessary between different data models, develop the mapping specification required for translating data between data model formats.
- (9) Document the information content of and relationships between data-in-transit and data service schema in a Data Resource Map. This map, like a common city road map, represents the data relationships between different data assets within the CE/mission area and enables the planning and understanding of how the information is used within the CE/mission area.

The end goal of the planning and organization process is a Data Resource Relationship Map that enables information engineers to understand what information is available from what data asset and how-and-why it moves between resources and consumers. It also provides an architectural planning view for developers to understand and implement data, data services, and messaging within a CE/mission area.

G.2.4.3 Data Model Guidance

There are many kinds of Data Model Guidance available. The purpose of the data model guidance includes:

- Education and Training in data modelling principles, techniques, and limitations;
- Consistency and uniformity of practice to ensure that tools and techniques are being understood and applied in the same way;
 - Inconsistency in the application of data modelling tools and techniques is a major reason why the use of data models to support application interoperability have had limited success;
- Engender, facilitate, and promote application interoperability through consistency of data model application.

Data model guidance includes:

- Standardized, reusable schema components for ubiquitous concepts (e.g., person, location, time) that can be incorporated into data models under development;
- Vocabularies and Taxonomies that establish common definitions of common terms;
- Guidelines:
 - Naming conventions; and
 - Modelling patterns, paradigms, styles;
- Training:
 - General data modelling principles and practices; and
 - Data mapping and translation;
- References:
 - Lists of metadata/schema registries, e.g., DoD DSE 2.0 [25]; and
 - Roster of data modelling SMEs;
- Tool recommendations and usage guidance.

In future versions of the AIA, this list will be expanded to include more categories of support resources and identify specific data model development support resources adopted and endorsed by the Army.

G.2.5 Interoperability Mapping, Translation, and Mediation Processes

Developing effective information sharing capabilities within a community and enabling interoperability among members of the community involves three (3) major areas of effort, each of which consists of a set of actions that must be completed. The three areas of effort are:

- Design: Charting, Modelling and Mapping;
- Operation: Exchange, Translation, and Mediation; and
- Maintenance: Troubleshooting and Tuning.

The set of actions required in each of these areas are described in the following sections.

Anticipated Information Sharing and Interoperability is based on the identification and makeup of a community (either an informal interoperability community or formal COI). Each member of the community is a pairing of a system/application (or data service) and a human POC/SME representing the system application. Data exchange between the systems mediated through the use of an IES, which defines the common data format for moving data from one system to another.

G.2.5.1 Design: Pathways, Modelling and Mapping

The design phase of creating an interoperability solution within a community assumes that the members of the community are known and participating in the process. The following actions must be completed before moving on to operation:

- Ensure that each member has a published, configuration-controlled schema that governs the data that is available from, or exposed by, their system/application or service;
- Ensure that the community has an IES with which to exchange data; the information represented by the IES must be a superset of the information that may be exchanged among the members (i.e., the IES must be able to “hold” all of the possible information that may be shared between/among members);
- Determine the information sharing “pathways” among members of the community; this consists of identifying and describing the anticipated “information sharing events” and the information shared between members; the purpose of this step is to establish the *information requirements* for interoperability and serve as basis for evaluating the completeness and fitness-for-use of the IES;
- Define the *mapping specifications* between each member schema and the IES; mapping specifications are directional, so one is needed from the member schema to the IES and from the IES to the member schema; the mapping specification shall be formal and detailed enough to unambiguously specify how data is transformed from one schema format to another; commercial tools are available for developing and documenting mapping specifications; and
- The information sharing pathways are used to evaluate the completeness and accuracy of the two-hop information sharing process: source member schema => IES format => target member schema; they are also used to troubleshoot information sharing errors.

The end result of this phase is a hub-and-spoke model of interoperability where the hub is the IES, the leaf-nodes are the member schemas, and the spokes are the mapping specifications.

G.2.5.2 Operation: Exchange, Translation, and Mediation

The hub-and-spoke model governs the actual information sharing process by controlling how data is exchanged among members of a community. The act of sharing information between members (i.e., exchanging data between systems) consists of the following steps:

- The data to be exchanged is extracted from the member’s data and forms the source dataset; this action can be initiated by the source member (in the case of a “push” information sharing event) or initiated by a service invocation by the consumer member (in the case of a “pull” information sharing event);
- The source dataset is translated from the member schema format to the mediating IES format; the translator software is derived directly from the mapping specification;
- The IES formatted dataset is translated to the target member schema formatted dataset; and
- The target dataset is received by the receiving member of the information sharing event; the data is then integrated into the data of the receiving member’s system.

The actual acts of data translation can occur according to one of four patterns:

- Translation from source dataset to IES dataset takes place on the source member's system; the IES dataset is transmitted over the network from source member to target member; translation from IES dataset to target dataset takes place on receiving member's system;
- Translation from source dataset to IES dataset takes place on the source member's system; the IES dataset is transmitted over the network from source member to a mediator agent on the network (e.g., ESB); translation from IES dataset to target dataset is performed by the mediator agent on the network; target dataset is transmitted to receiving member;
- Source dataset is transmitted over the network from source member to a mediator agent on the network (e.g., ESB); translation from source dataset to IES dataset is performed by the mediator agent on the network; IES dataset is transmitted to receiving member; translation from IES dataset to target dataset takes place on receiving member's system; and
- Source dataset is transmitted over the network from source member to a mediator agent on the network (e.g., ESB); translation from source dataset to IES dataset is performed by the mediator agent on the network; translation from IES dataset to target dataset is performed by the mediator agent on the network; target dataset is transmitted to receiving member;

G.2.5.3 Maintenance: Troubleshooting and Tuning

Information sharing acts will reveal errors in the mapping process due to causes such as misinterpretation of the IES or ambiguous representation of information in member schemas. When an information sharing error occurs, the following steps are used to analyze and correct the causes of the error:

- From the error discovered on the target system, trace the translation process backwards from the target member schema to the IES schema using the mapping specification;
- Analogously, trace the translation process backward from the IES schema for the source member schema using the mapping specification;
- Sources of error include (but are not limited to):
 - Incongruent interpretations of IES elements;
 - Semantic incompatibilities between IES and member schemas (e.g., information that can be represented by a member schema cannot be represented by the IES); and
 - Unforeseen semantic subtleties (e.g., does a field "meal cost" include tip?);
- Based on the cause discovered, correct the mapping specification, member schema(s), and/or IESs; and
- Regenerate translation code based on mapping specification updates.

G.3 Service Planning Processes

G.3.1 Data Service Development Process

The development of the DSL-A [18] data service interface specifications included a description of the data service development process. The overall, general process is illustrated in Figure 17 and described in Table 14. Table 14: Data Service Development Process Description

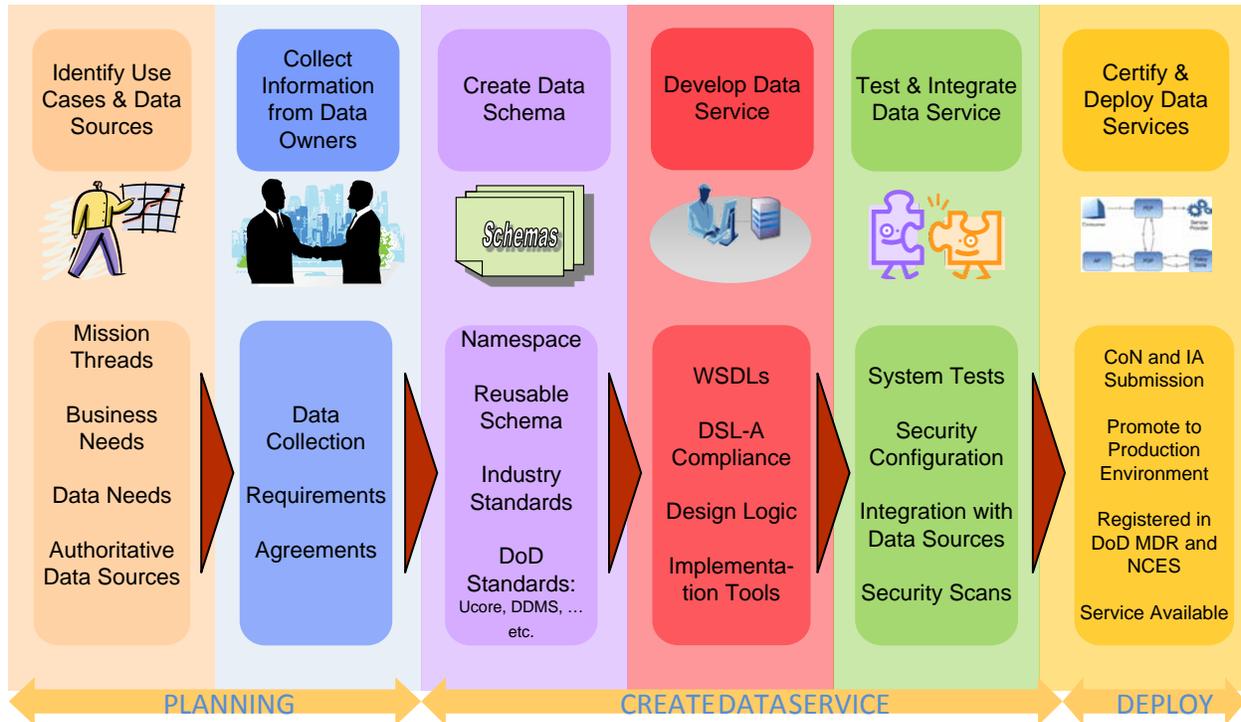


Figure 17: Data Service Development Process

Table 14: Data Service Development Process Description

Process Phase	Description
Identify Use Cases & Data Assets	Use Cases establish the business reasons why the data service(s) is needed and identify the information requirements that must be met by the service. The data assets to which the data service provide access are identified based on the requirements specified by/in the Use Cases. During this phase, the Service Portfolio Management processes ensure that the prospective service (a) does not already exist, and (b) fits logically within the capabilities of the portfolio. ADSs may either supply data to the service (DSE 2.0 should be consulted to search for data that meets needs), or the data service provides access to a candidate ADS. The DSE 2.0 should be consulted to search for existing services that may supply required data.
Collect Information from Data Owners	Data owners are engaged to obtain information about the data asset (e.g., the data asset schema) and to negotiate and agree to Data Use and Web Service Accessibility agreements.
Create Data Schema	Schemas are created to specify and govern the data available through the data service interface. Army namespace standards will be applied/evaluated, as well as use of standardized schemas. Schemas will be registered with the DSE 2.0.
Develop Data Service	The Data Service will be developed using the IES and WSDLs will be created or adopted from standardized service interfaces specifications. Support tools (e.g., CDSF) may be used to create and test the data service. The data service will be tested for compliance to Army service requirements (e.g., security).

Process Phase	Description
Test & Integrate Data Service	The developed service will undergo final testing processes (e.g., NETCOM, STIGs) and security options will be configured (including security handlers and ABAC policies) before deployment.
Certify & Deploy Data Service	As part of data service deployment, the data service will be submitted to CoN and IA for certification and be registered with DoD service registries (e.g., DSE 2.0).

G.4 Migrating Data to a Cloud Computing Environment

Deciding to move enterprise software applications to CCE involves a tradeoff analysis between the benefits offered by a cloud (e.g., low start-up costs, to pay-as-you-go resource usage, unlimited expandability) and the risks/limitations associated with running software in a cloud (e.g., security, bandwidth). Moving enterprise data to a cloud computing environment is a distinct and different decision-making process when compared to moving applications to the cloud - it involves a separate set of benefits and risks. The chief benefit of moving data to a cloud is global shared data access; the chief risk is security of the data. A framework is presented that describes a structured set of factors that an enterprise should consider when deciding whether or not to move data to a cloud environment. The framework may be also used as an evaluation for evaluating a cloud data deployment. The end-goal of the framework application is secure and *trusted* cloud data. The approach could be described as a strategy for assessing the “cloud-ability” of enterprise data.

G.4.1 Introduction

The purpose of this section is to define a framework of factors to be considered by an enterprise when deploying data to a CCE or assessing the level of implementation maturity of data deployed in a CCE.

Data management practices or data strategy are not addressed, though the subjects addressed here may be considered an element of both data management practices and strategy.

The objectives of this section are:

- Describe the motivations and reasons for moving data to a CCE;
- Identify and describe the significant factors that must be considered by an enterprise when moving data to and maintaining data in a CCE; and
- Provide a check-list of “things-to-do” before, during, and at the end of deploying data to a CCE.

See the *Army Data Framework - Data Aspects of Cloud Computing* [37] for additional guidance on the implementation of cloud computer solutions.

G.4.2 Cloud Data Deployment

G.4.2.1 The Fives Axes of Cloud Data Deployment

There are five (5) relatively independent axes of factors that need to be considered when moving data to a CCE:

- Use: Why move data to a cloud?
- Security
- Legal
- Technical Design
- Implementation, Operation, and Maintenance

G.4.2.2 Consideration #1: Use - Why move data to cloud?

The first consideration for moving data to the cloud is: Why? The following are possible business drivers that lead to a decision to move data to a cloud:

- Strategic Direction of Enterprise: The long term business strategy of the enterprise is conducive to cloud computing;
- Efficiency / cost savings: Buy versus Build versus Re-use:
 - Scalability / Expandability;
 - failsafe storage; security from loss/corruption;
 - leverage cloud infrastructure expertise/economies of scale; for example, an enterprise may not want to invest in security and would trust the cloud hosting vendor to be secure;
 - Resource Reuse; and
 - High Availability;
- Business Process Support: Example: Business Intelligence is a particular application that draws on global access, data integration, and collaboration:
 - Global access by dispersed clients;
 - Commoditized data; Data Marketplace;
 - Data Integration; and
 - Collaboration; process/application interoperability;
- Who is served? Stakeholders, applications - business objectives for deploying data to cloud; and
- Technology experimentation and research.

If one of these reasons/objectives is not driving the consideration of cloud technology, chances are cloud computing is not a good choice. Moving data or applications to the cloud just because it is the latest new technology or “everyone is doing it” is not a sound reason for doing so.

G.4.2.3 Consideration #2: Security

Security and Legal (see next section) considerations are the most significant factors that need to be addressed prior to moving data to a CCE. Items that need to be considered include:

- Cloud Service Consumer Precautions:
 - Sensitivity/Value of Data; need to know to evaluation suitability of cloud service provider security capabilities;

- Cloud Host/Provider Precautions:
 - Security (includes Federal Information Security Management Act (FISMA) certification):
 - Access control (theft, compromise):
 - External Intruders / Hackers; and
 - Internal spies/malcontents;
 - Physical Security:
 - Physical Location of data (Includes export control consideration);
 - replication/archive sites; and
 - "Depth" of cloud (does the cloud service use other cloud service providers?);
 - Monitoring/Reporting planning, scheduling, and techniques;
 - Technical Precaution Techniques:
 - Authentication;
 - Encryption;
 - Monitoring; and
 - Auditing;
 - Tenant Isolation/Insulation; security precautions dealing with multiple tenants in single cloud environment.
 - Data Integrity (Backup, archive, replication plan); and
 - Client Inspections/Audits – are they allowed/encouraged?

See Sections 3.1.4 and 3.1.5 of the *Army Data Framework (ADF): Data Aspects of Cloud Computing* [37] for further information on cloud security.

G.4.2.4 Consideration #3: Legal

Because a cloud service is sometimes provided by a third party, the legal contract between the consumer and the provider must make clear the following considerations:

- Liability (Related to Security) – who is responsible for security breaches or data loss?
- Privacy (Related to Security) – how is consumer data privacy protected?
- Control/ownership of data; and
- Export Control.

Service Level Agreements (SLAs) and/or legal contracts define/clarify these items.

G.4.2.5 Consideration #4: Technical Design

The following technical design decisions must be considered and a choice made. The choices will depend on the reason for putting data into a CCE (see consideration #1).

- Deployment Model: Private, Public, Community, or Hybrid Cloud;
- Service Model: DaaS, DBaaS, SaaS, PaaS, or IaaS; and
- Data Architecture
 - Number of distinct data assets under single governance entity;

Information Integration and redundancy – how are the distinct data assets integrated or related to one another?

- Data storage paradigm, e.g.,:
 - Relational;
 - Key-value;
 - Dimensional;

- Data access paradigm:
 - Service;
 - Unique/proprietary API;
 - Connector/adapter;
- Physical data processing location:
 - Data Size versus Bandwidth Requirements – heavy duty analytics on Big Data cannot be done if processing is not “near” or within the same environment as the data due to bandwidths limitations;
 - Response time – bandwidth limitations, connectivity challenges (e.g., disadvantaged networks), and physical distance may make response time unacceptable;
- Data Quality – not unique to Cloud, needs to be done regardless of where data is; and
- Vendor Lock-in / Portability (See Termination Migration Plan)

G.4.2.6 Consideration #5: Implementation, Operation, Maintenance

Once data is deployed to the cloud and is in use, how will it be managed and maintained? This consideration is included in the decision-to-move-data-to-cloud process because there may be aspects of implementation, operation, and maintenance that impacts the design and/or technical design.

- Lifecycle Plan: What is the plan for the data for next year? Five years? For retirement/migration?
 - Termination Migration Plan – when it is time to move data off the cloud, how will that be done? Plans for erasing/zeroing out residual data?
- Operations beyond SCRUD: What kind of data services are needed beyond SCRUD? This is related to business use of data;
- Testing / Pilots; and
- Scheduled Audits (or any scheduled activity) are regular client activities to test, assess, or evaluate data quality, security, or other aspects of their data in the cloud. It is not advisable to just "put data out there" - need to "check the health" of the cloud data on a regular basis.

G.4.2.7 Summary

Many of the technical decisions that must be made before moving data to a CCE are the same as any data-centric implementation (e.g., moving data to a datacenter). Some considerations and decisions, however, are unique to cloud environments and are critical to the decision to move data to a cloud.