February 14, 2025

AIRC 2024-2025 Innovation Project

Enabling Data Management of Intellectual Property License Rights for the Department of Defense

Benjamin McMartin, Esq.
John (Jerry) G. McGinn, PhD
Lloyd Everhart
John Howard Davis
George Mason University

Joanne Herring, Program Management and Implementation Lead

OUSD (A&S) / ODASD (ADA)

Cleared for Public Release.

Abstract

The Department of Defense (DoD) is seeking to enhance its Data Management practices for Intellectual Property (IP) License Rights by researching commercial best practices, standards, and technologies. This study identifies the role of data structuring, cleansing, tagging, ontologies, taxonomies, Natural Language Processing (NLP), and Artificial Intelligence (AI) models in optimizing data management. Current DoD applications, including Joint All Domain Operations, Senior Leader Decision Support, and Business Analytics, utilize AI-driven platforms such as JADC2, ADA, and Advana to facilitate structured data retrieval and decision-making. The study also highlights commercial best practices, such as metadata governance, compliance with ISO standards, and adherence to the FAIR principles for improved data usability.

Challenges to adoption include the complexities of data valuation, enterprise-wide governance, and security concerns. To overcome these obstacles, the report recommends leveraging the Enterprise Visibility and Management of Operating and Support Cost (EVAMOSC) model, establishing secure data lakes, implementing AI-powered data transformation, and developing intuitive dashboards for improved decision-making. Moving forward, the DoD should focus on identifying data sources, evaluating storage solutions, defining security protocols, and formulating an IP-specific data management policy. By integrating commercial methodologies with AI-driven analytics, the DoD can achieve a more secure, interoperable, and efficient data management framework.

Executive Summary

The Department of Defense (DoD) seeks to improve Data Management practices for Intellectual Property (IP) License Rights by researching commercial best practices, standards, and technologies. This involves mapping the license rights chain, identifying challenges in adoption, and exploring potential solutions.

Key Findings

1. **Data Management Applications:**
    - Effective data management involves structuring, cleansing, tagging, ontologies, taxonomies, NLP, and AI models.
    - DoD applies data management in Joint All Domain Operations, Senior Leader Decision Support, and Business Analytics.
    - Existing platforms such as JADC2, ADA, and Advana use AI/ML for data structuring, retrieval, and decision-making.
2. **Commercial Best Practices & Standards:**
    - Industry best practices emphasize metadata management, compliance with standards (ISO 27001, ISO 8000, DAMA-DMBOK), and governance frameworks.
    - FAIR principles ensure data is Findable, Accessible, Interoperable, and Reusable.
    - Standards like CRISP-DM guide structured data mining processes.
3. **Challenges to Adoption:**
    - Data valuation complexities and inconsistent standards hinder implementation.
    - Collaboration across multiple stakeholders and ensuring enterprise-wide governance is challenging.
    - Security concerns require automation, encrypted storage, and controlled access.
4. **Proposed Implementation Strategy:**
    - Utilize the Enterprise Visibility and Management of Operating and Support Cost (EVAMOSC) model as a framework.
    - Establish secure, scalable data lakes and authoritative databases for structured data management.
    - Implement AI-driven data transformation processes to automate data cleansing and compliance checks.

- Deploy user-friendly dashboards and business intelligence applications to facilitate decision-making.

Recommended Actions:

1. Research and identify data sources and characteristics.
2. Evaluate data storage options, including data lakes and structured databases.
3. Define data characterization, handling, and security requirements.
4. Develop an IP-specific data management policy aligned with commercial best practices.

By leveraging commercial standards and AI-driven data management techniques, DoD can enhance its IP data governance, security, and interoperability while addressing challenges in implementation and adoption.

Research Team

Name	Organization	Labor Category
Benjamin McMartin	George Mason University	Co-PI
John (Jerry) G. McGinn	George Mason University	Co-PI
Lloyd Everhart	George Mason University	Researcher
John Howard Davis	George Mason University	Graduate Research Assistant

Table of Contents

List of Figures

Acronyms and Abbreviations

3NF	Third Normal Form
ABAC	Attribute-based Access Controls
ADA	Acquisition Data and Analytics
AES-256	Advanced Encryption Standard-256
AI	Artificial Intelligence
AI/ML	Artificial Intelligence/Machine Learning
API	Application Programming Interface
ARMA	Association of Records Managers and Administrators
CAC	Common Access Card
CD	Compact Disc
CDAO	Chief Digital and Artificial
CDO	Chief Data Officer
CIA	Central Intelligence Agency
CIO	Chief Information Officer
CRISP-DM	Cross-Industry Standard Process for Data Mining
CSF	Cybersecurity Framework
DAMA	Data Management Association
DAMA-DMBOK	Data Management Body of Knowledge
DARPA	Defense Advance Research Projects Agency
DCAM	Data Management Capability Assessment Model
DFARS	Defense Federal Acquisition Regulation Supplement
DMCA	Digital Millennium Copyright Act
DoD	Department of Defense
DOE	Department of Energy
DPRIVE	Data Protection in Virtual Environments
DRM	Digital Rights Management
DVD	Digital Video Disc
EFF	Electronic Frontier Foundation
ETL	Extract Transform and Load
EVAMOSC	Enterprise Visibility and Management of Operating and Support Cost
FAIR	Findability Accessibility Interoperability Reusability
FAR	Federal Acquisition Regulation
FedRAMP	Federal Risk and Authorization Management Program
FHE	Fully Homomorphic Encryption
IEC	International Electrotechnical Commission

IP	Intellectual Property
ISO	International Organization for Standardization
IT	Information Technology
JADC2	Joint All Domain Command and Control
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
MFA	Multi-Factor Authentication
MFT	Managed File Transfer
ML	Machine Learning
NASA	National Aeronautics and Space Administration
NATO	North Atlantic Treaty Organization
NDS	National Defense Strategy
NER	Named Entity Recognition
NIST	National Institute of Standards and Technology
NLP	Natural Language Processing
NSF	National Science Foundation
ODRL	Open Digital Rights Language
OTA	Other Transaction Agreement
PII	Personally Identifiable Information
PIV	Personal Identity Verification
POS	Part of Speech
R&D	Research and Development
RBAC	Role-based Access Control
RDBMS	Relational Database Management System
SME	Subject Matter Experts
SNLR	Specifically Negotiated License Rights
SQL	Structured Query Language
SSO	Single Sign-On
TLS	Transport Layer Security
USPTO	United States Patent and Trademark Office
UUID	Universally Unique Identifier
VAULTIS	Visible Accessible Understandable Linked Trusted Interoperable Secure
XML	Extensible Markup Language

1. Introduction

Research commercial best practices, standards, and technologies available for adoption to support data management of the Department of Defense (DoD) intellectual property license rights including mapping the chain of license rights from solicitation documents, proposal documents and assertions, contract instruments, and deliverables, identifying potential challenges to DoD's adoption and use of commercial best practices, standards, and technologies, and identifying potential paths to adoption of commercial best practices, standards, and technologies which address the challenges identified.

Organization of the Report

Chapter 1 this introduction describes the objectives of the research effort.

Chapter 2 answers the question what is Data Management?

Chapter 3 answers the question what are the practical applications of Data Management?

Chapter 4 answers the question which practical applications of Data Management are potentially relevant to DoD's management of Intellectual Property rights?

Chapter 5 answers the question what are the commercial best practices, standards, and solutions available for Data Management applications relevant to DoD IP management?

Chapter 6 answers the questions can the best practices, standards, and solutions available for Data Management applications relevant to DoD IP management be adopted? What are the challenges to adoption? What actions would be required to overcome the challenges of Data Management adoption within the DoD?

Chapter 7 Recommends pathways to implementation of relevant commercial Data Management best practices, standards, and solutions.

2. What is Data Management?

Data management can be defined as the process of ingesting, storing, organizing, and maintaining data created and collected by an organization [Stedman, 2024]. Organizations can then efficiently use that data to run operations and motivate business decisions [University of Phoenix, N.D.]. Examples of data management processes that perform these activities include data structuring, cleansing, tagging, ontologies, taxonomies, natural language processing (NLP), and use of artificial intelligence models for data curation and retrieval.

3. What are the practical applications of Data Management?

Data management has various practical applications that require data structuring, cleansing, tagging, ontologies, taxonomies, natural language processing, and use of artificial intelligence or machine learning models.

Data structuring refers to the process of collecting both structured and unstructured data before converting it into a usable form [Tableau, N.D.]. These forms may include arrays, trees, queues, or stacks [Tableau, N.D.], which can be applied for various practical uses. For example, arrays are used for storing data for computations, performing image processing, and record management [CCS Learning, 2024]. In addition, trees are used to develop decision-based algorithms for machine learning and perform indexing [CCS Learning, 2024]. Furthermore, queues can handle website traffic and interruptions in operating systems, as well as serve requests on a shared resource, while stacks can monitor previously visited sites [CCS Learning, 2024]. In general, these forms can also be used to create graphics like ranked lists, charts or graphs, process progression and degression, and many other tools for decision making [Tableau, N.D.].

Data cleansing is necessary to edit data to remove incorrect, duplicate, or corrupted pieces from a dataset [Manage Engine, N.D.]. This process works by removing irrelevant data, fixing formatting errors, removing incorrect outliers if necessary, managing missing data either by deleting it, reworking the algorithm to ignore it, or estimating it, as well as validating the cleansed data [Manage Engine, N.D.]. Data cleansing is applicable to data management for customer, sales, financial, and human resource data [AltexSoft, 2021]. For example, data cleansing is necessary to ensure accurate customer data including name and contact information, as well as accurate human resource data with employee information [AltexSoft, 2021]. Furthermore, sales information like product descriptions, price, and value are kept organized and correct using this process [AltexSoft, 2021]. Additionally, financial records including revenue, costs, and taxes are kept accurate for compliance using cleansing [AltexSoft, 2021].

Data should be tagged using labels or metadata to facilitate simpler searching [Merced, 2024]. These labels often include creator's name, date, department, file format, or other relevant information [Merced, 2024]. Tags can be applied to text with named entity recognition (NER) or part of speech (POS), or videos with 2D bounding and semantic segmentation [Merced, 2024]. Additionally, to use data tagging, organizations can follow models like the hierarchical, flat, segment, or jargon models [Merced, 2024]. Applications of data tagging include to help facilitate easier identification, categorization, and stronger data security [Merced, 2024], as well as to tag text for natural language processing [Shane, 2017]. Examples of these applications include identifying spam and non-spam messages, classifying documents, and natural language processing uses like explaining text meaning and word relationships for machine translation, designing conversations for chatbots, and identifying consumer feelings from reviews [Shane, 2017].

The data must also be organized using a system, one of which is ontologies. Ontologies enhance domain descriptions using classes, relationships, and instances to support reasoning about entities within a domain [Amplitude, N.D.]. In these descriptions, classes refer to fundamental categories of objects, relationships work between classes, and instances are the individual data points or objects [Amplitude, N.D.]. This framework makes it easier to understand the data from another source, which promotes data interoperability, integration, and discovery and governance [Amplitude, N.D.]. Examples of ontologies in practical use include data integration from multiple disparate systems to identify system failures for NASA, using the Semantic Sensor Network for environmental monitoring, and for market volatility and risk management for finance [The Data Governor, 2023].

Organizations may also use taxonomies to categorize a domain's objects hierarchically [Amplitude, N.D.] by creating groups based on the objects' characteristics, attributes, and relationships [Wharton, 2024]. These groups are then organized into categories and subcategories [Wharton, 2024]. Examples of practical applications of taxonomies include organizing a product catalog [Marin, 2023], books in a library [Marin, 2023], or business data [DoD JADC2, 2022] into categories and subcategories for more efficient access and use. For example, business data, including customer, market trend, financial and performance, and competitor data need to be organized to allow professionals to easily access the relevant metrics for decision making [DoD JADC2, 2022]. Therefore, data taxonomies have the usual benefits for a data organization method, including a consistent form of organization, easier data access, improved data quality, efficient analysis, and easy compliance [Wharton, 2024].

There are faster ways to facilitate data management, though, like using AI models and NLP to automate the processes. Natural language processing is the process of teaching a computer to understand, interpret, and use human language [Kosinski, 2024]. NLP connects human and computer languages to help facilitate more efficient business processes [Kosinski, 2024]. For example, natural language processing can process and filter large quantities of unstructured data for easier data analysis [Kosinski, 2024]. During this process, NLP identifies patterns, extracts data without human error, performs data analysis, and more [Kosinski, 2024].

Artificial intelligence models can also be used to automate data management processes including data collection, cleaning, analysis, and security [OSD, N.D.]. These steps aim to improve data discovery, quality, accessibility, and security [OSD, N.D.]. For example, AI models can use data discovery to identify all data that an organization possesses, which can be useful for easy access and prevention of data breaches of this previously hidden data [OSD, N.D.]. Additionally, AI models can use data cleansing to fix errors and promote data quality, as well as integrate data to improve accessibility [OSD, N.D.]. Examples of AI models performing these tasks include scanning network devices to discover data and index it, performing validation checks and flagging or fixing errors, automatically detecting relationships between datasets to integrate them, and automatically detecting personally identifiable information (PII) or other private information and blocking unauthorized use [OSD, N.D.].

4. Which practical applications of Data Management are potentially relevant to the DoD's management of Intellectual Property rights?

The DoD applies data management towards three main focus areas: Joint All Domain Operations, Senior Leader Decision Support, and Business Analytics [Department of Defense, 2020]. Within each of these areas, the DoD uses various forms of data management, as mentioned above, for many different applications.

In Joint All Domain Operations, the DoD currently uses a Joint All Domain Command and Control (JADC2) system that creates a centralized data platform across branches [DoD Financial, 2022]. The JADC2 applies data structuring, tagging, and artificial intelligence technology to streamline data access and analysis for all domains [DoD Financial, 2022]. By having all data across all branches on one system, the data must all be structured, or formatted, correctly to facilitate easier access [DoD Financial, 2022]. This can include a minimum metadata tagging criteria [DoD Financial, 2022]. Additionally, JADC2 uses artificial intelligence to manage data with automatic machine-to-machine transactions that extract, collect, and process large quantities of data inputted by the sensing infrastructure [DoD Financial, 2022]. These applications of data structuring, tagging, and AI use help Joint Force Commanders better lead the Joint Force across all warfighting domains through quick, accurate decision making, as well as help to defeat adversaries with secure information [DoD Financial, 2022]

For Senior Leader Decision Support, DoD senior leaders, including the Deputy Secretary, are requesting the development of metrics to inform their management decisions, such as how to implement the National Defense Strategy (NDS) [Department of Defense, 2020]. To accomplish this, the DoD will apply data structuring, NLP, and machine learning technologies for data analysis through the Office of Acquisition Data and Analytics (ADA) [OSD Data Analytics, N.D.]. For data structuring, the ADA uses tree-based modeling to describe, categorize, and generalize the data of, for example, total and procurement costs [Stedman, 2024]. Additionally, NLP is used with text mining to analyze patterns in frequency of words to identify topics of concern [Stedman, 2024]. Furthermore, the ADA uses machine learning for applications in classification, like identity fraud detection or image classification; regression with market forecasting; for real-time decisions, robot navigation, and game AI; for dimensionality reduction with big data visualization; and in clustering with recommender systems and customer segmentation [Stedman, 2024]. Furthermore, machine learning is used to create visualizations of data including schedules and cost growth [Stedman, 2024]. These applications of data structuring, NLP, and machine learning technologies will assist the DoD's transition to current, interactive data instead of slides to motivate senior leaders' decision making [Department of Defense, 2020].

For Business Analytics, DOD is working on an effort to ingest, analyze, and display business data [Department of Defense, 2020]. This business data would include budget, procurement, inventory, logistics, and personnel information [Department of Defense, 2020]. To complete this effort, the DoD is applying the platform Advana [Lin, 2021], which uses data structuring and AI/ML technologies to manage business data. For example, Advana structures data in usable formats for analysis through the Qlik feature, which creates visualizations for decisions [DoD Chief Digital, 2024]. Additionally, Advana uses machine learning algorithms against data for query and analysis [Sokolowski, 2023]. These applications of data structuring and AI/ML technologies in Advana allow the DoD to use a common data platform for all business analytics [Department of Defense, 2020]. Additionally, this effort can give insights to improve future data efforts through informing the data governance community on best practices for policies on authoritative data sources, consistent metadata labeling, standard taxonomies, data provenance, and interfaces [Department of Defense, 2020].

These practical applications of data management, including data structuring, tagging, NLP, and AI/ML technologies, through JADC2, ADA work for the NDS, and Advana facilitate better decision making for the DoD.

5. What are the commerical best practices, standards, and solutions available for Data Management applications relevant to DoD IP management?

Many commercial best practices exist that the DoD can draw from to improve their data management practices. Commercial agencies that are building services for data discovery and establishing guidelines to manage costs, sensitive data protection, and legal compliance are especially good examples [Holloway, 2022]. The DoD can learn from their commercial experiences and improve their own data management activities by bringing in experts, contributing additional funds to innovation and special projects, attending conferences and events, training their workforce on cloud, and automating data access for analytics [Holloway, 2022]. The DoD should bring in experts, from military and civilian sources and through public and private partnerships, to introduce new ideas and innovation efforts [Holloway, 2022]. These experts can help improve data fluency by using fresh techniques to address skill gaps for the DoD workforce [Holloway, 2022]. The DoD should also contribute funds to innovation and special projects for prototype and innovative approach development [Holloway, 2022]. DoD agencies should use Other Transaction Agreements (OTAs) and other acquisition or funding programs to make this contribution [Holloway, 2022]. Furthermore, the DoD should encourage participation in conferences and events that highlight successful applications [Holloway, 2022]. These conferences can provide insights on technologies that improve data readiness and analytics, data governance and compliance, data tools, and data-based government decision-making [Holloway, 2022]. Additionally, the DoD should train their workforce in navigating modern cloud infrastructure to automate the data discovery process using AI/ML technologies. As the DoD transitions from a single to a multi-cloud provider with the Joint Warfighter Cloud Capability and the CIA's Commercial Cloud Enterprise contract, pressure will increase on IT and security teams who will require more training [Holloway, 2022]. The DoD also needs to automate data access for analytics at the data layer that focuses on zero trust [Holloway, 2022]. The zero trust, or the automated monitoring layer, leaves a security gap, meaning that the DoD needs to use an automated attribute-based access control model to remedy this [Holloway, 2022].

Furthermore, the DoD should use data management systems in the same ways as other commercial users, including: to secure data storage on hardware and software, to abide by compliance and regulatory standards, for data creation and accessibility from a diverse source pool, to promote data availability in spite of network strain or failure, and to incorporate data into applications and analytics programs [University of Phoenix, N.D.]. These applications of data management rely on best practices like long and short-term storage, back-ups, using file naming, and proper note-taking on data management systems [University of Phoenix, N.D.], as well as identifying business goals, emphasizing data quality and security, and allowing authorized users easy access on the organization level [Tableau, 2024]. For storage, short- term storage is useful, and can be facilitated using computers, shared servers, and cloud storage [University of California, N.D.]. Long-term systems, like preservation systems [University of California, N.D.], are useful to keep data secure but accessible when needed [University of Phoenix, N.D.]. Other important storage conventions, like only using flash drives for transfer, not storage, and storing copies of data in accessible and stable formats, as well as keeping the original copy, should be practiced [University of California, N.D.]. Backing-up data on local and cloud servers maintains the security and accessibility of the data in case of a server failure [University of Phoenix, N.D.]. The Rule of 3 is important for back-ups, as it recommends that organizations keep 2 copies onsite and 1 offsite, while performing regular back-ups to maintain the data [University of California, N.D.]. Data back-ups also help with preservation, which keeps data secure and accessible for the future [University of California, N.D.]. Organizations can facilitate preservation by identifying data of potential value and preserving it and its derivatives, as well as any code and corresponding software programs that help with transforming, analyzing, and understanding the data [University of California,

N.D.]. Additionally, file naming, especially with relevant identifying information and including metadata with its schemes and standards [University of California, N.D.], is important for organization and easy access, while note-taking allows this data to be put into a relevant context [University of Phoenix, N.D.]. On the organization level, setting goals determines what type of process should be used for data management to manage the most relevant data to an organization's aimed business decisions [Tableau, 2024]. Additionally, data quality and security is important and can be implemented by training team members on proper data handling [Tableau, 2024], inputting [Tableau, 2024], and clean-up [University of California, N.D.]. Furthermore, allowing the authorized people easy access to necessary data improved the efficiency of business operations [Tableau, 2024].

Important standards relevant to DoD IP and data management include the Procurement Data Standard, which was developed under guidance from the Office of Under Secretary of Defense for Acquisition and Sustainment's' Office of Defense Procurement and Acquisition Policy office, to standardize data creation, translation, processing, and sharing for procurement [Hero, 2024]. This standard outlines the minimum requirements for contract writing system output with the aim of improving visibility and accuracy of contract data, interoperability for acquisition systems, and standardization of the procure-to-pay process [Hero, 2024].

Additional existing standards stem from organizations like the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) [Desai, 2023]. These organizations work together as the authority to establish global standards for IP management, data management, and security [Desai, 2023]. Standards relevant to data management specifically are often also related to data governance, as data governance standards control the actions taken in data management, as well as decide who completes the actions [Desai, 2023]. Data governance is regulated by ISO and IEC standards, which are established to quantify data management through measures of quantity, weight, extent, value, or quality [Desai, 2023]. Important ISO Standards for data governance include ISO 27001 and ISO 8000, with the former focusing on information security management systems [Desai, 2023], including how data is protected, and the latter on data quality management [Desai, 2023]. For example, ISO 8000 includes syntax, origins, completion, accuracy, and certification [Desai, 2023] to ensure that data is accurate, consistent, and usable. Other relevant standards include ISO/IEC 38505-1:2017, ISO 22745, ISO 3166, ISO/IEC 11179, and ISO/IEC 27701:2019 [Desai, 2023]. ISO/IEC 38505-1:2017 is a standard for data governance within an organization, focusing on providing a model for evaluating, directing, and monitoring their data handling and usage [Desai, 2023]. This standard also emphasizes IT led data governance [Desai, 2023], including data and information management. Additionally, ISO 22745 specifies data requirements for master data, including syntax, encoding, and portability based on the NATO Codification System [Desai, 2023]. Furthermore, ISO 3166 defines codes for country names, facilitating easier access to reference data with international and national extensions [Desai, 2023]. ISO/IEC 11179 details metadata registry standards to make data comprehensible and sharable [Desai, 2023]. Finally, ISO/IEC 27701:2019 extends the aforementioned ISO 27001 with a data privacy extension to facilitate compliance with data privacy regulations [Desai, 2023].

Other standards like the DAMA-DMBOK (Data Management Body of Knowledge) from the Data Management Association (DAMA) provide a comprehensive guide for data management [LinkedIn, N.D.].

Furthermore, standards like Findability, Accessibility, Interoperability, and Reusability (FAIR) principles, which were developed for digital assets and play an important role in data management [Go Fair, 2016]. These standards improve data findability by making metadata and data easy to find for humans and computers, with computers requiring machine-readable metadata for improved efficiency [Go Fair, 2016]. Accessibility requires users to know how to access data they found, including authentication and authorization [Go Fair, 2016]. The

FAIR principles improve interoperability by integrating data with other data to interoperate with other applications for analysis, storage, and processing [Go Fair, 2016]. Furthermore, they increase reusability by ensuring that metadata and data are properly described so that they can be replicated or combined in other settings [Go Fair, 2016].

In addition, more specific standards for data management exist in domains like data mining, where the Cross-Industry Standard Process for Data Mining, or CRISP-DM, governs operations [Data Science, 2024]. The CRISP-DM is built off of six phases, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment [Data Science, 2024]. Business understanding focuses on identifying goals, assessing the situation, determining specific data mining objectives, and creating a project plan [Data Science, 2024]. Data understanding emphasizes collecting initial data, before describing, exploring, and verifying it [Data Science, 2024]. Furthermore, data preparation works on selecting, cleaning, constructing, integrating, and formatting the data, while modeling requires users to select a technique, test the design, build the model, and finally assess it [Data Science, 2024]. For evaluation, users should evaluate their results, review the overall process, and determine their next steps [Data Science, 2024]. Deployment requires planning, monitoring and maintenance, producing a final report, and reflecting on the project [Data Science, 2024].

Additional standards come from organizations like the NIST, who provide an important data management manual on data protection in their cybersecurity framework (CSF) [National Institute, 2024]. The second edition of this guide provides standards for industry, government, and others on managing cybersecurity [National Institute, 2024]. The CSF gives outcomes that allow organizations to understand, assess, prioritize, and communicate cybersecurity measures [National Institute, 2024]. Additionally, it recommends other resources with further advice on reaching these outcomes [National Institute, 2024].

A solution that addresses the DoD's data management application in online data protection is the Data Protection in Virtual Environments (DPRIVE) solution [DARPA, 2024]. This solution is needed because current solutions in this field have adequate data protection during transmission, at rest, and in storage, but require decrypting data for computations. [DARPA, 2024] This weakens data security, so the Defense Advanced Research Projects Agency (DARPA) developed the DPRIVE solution, which uses fully homomorphic encryption (FHE), to enable computation on encrypted data [DARPA, 2024]. Although FHE is successful at running calculations on encrypted data, it requires a long time to perform simple tasks, making it unsuitable to use as a standard processing hardware [DARPA, 2024]. The DPRIVE solution aims to enable FHE computations to run at the same pace as unencrypted computations, therefore strengthening data security for the DoD and other commercial organizations [DARPA, 2024]. This is achieved by developing a hardware accelerator that reduces processing overhead, therefore improving the speed of FHE calculations [DARPA, 2024]. In particular, DPRIVE will do this by building new methods of memory management, flexible data structures and programming models, and formal verification methods to improve computation efficiency and accuracy [DARPA, 2024]. The DPRIVE system will reduce run time for current FHE computations by many magnitudes in line with performance on unencrypted data [DARPA, 2024].

Another solution relevant to the DoD's data management applications for IP management is Google Cloud's Apigee [Brown, 2024]. This solution works within a previous system, Advana, to further develop the DoD's data sharing capabilities [Brown, 2024]. The previous solution, Advana, was successful at fulfilling most of the DoD's requirements in scale, security, and stability, yet it struggled in allowing the right DoD personnel to have easy access to the right data at the right time [Brown, 2024]. This is because Advana did not address data silos [Brown, 2024], or collections of data controlled by one department and isolated from the rest of an organization

[Stedman, 2023], which prevented quick and effective access, use, searching, and sharing of data [Brown, 2024]. Therefore, the Chief Digital and Artificial Intelligence Office (CDAO) began utilizing Google Cloud's Apigee program, in combination with Advana, to enhance their data use capabilities by creating a decentralized data mesh using an application programming interface (API) [Brown, 2024]. Apigee allows the DoD to manage a high volume of APIs in Advana, which leads to improved data flow, streamlined processes, and increased communications [Brown, 2024].

Furthermore, EDM Council's Data Management Capability Assessment Model (DCAM) framework proves to be a useful tool for data management applications [EDM Council, N.D.]. This model uses four levels and eight components to foster data governance for organizations [EDM Council, N.D.]. These levels include Foundation, Execution, Collaboration, and Application [EDM Council, N.D.]. On the foundational level, the model focuses on building a data strategy and business case, as well as implementing a data management program with funding [EDM Council, N.D.]. For execution, this framework emphasizes improving business and data architecture, data and technology architecture, data quality management, and data governance [EDM Council, N.D.]. Additionally, collaboration refers to creating a data control environment, and application mentions analytics management [EDM Council, N.D.]. After implementing this framework, DCAM is scored based on engagement, process, and evidence [EDM Council, N.D.].

6. Can the best practices, standards, and solutions available for Data Management applications relevant to DoD IP management be adopted? What are the challenges to adoption? What actions would be required to overcome the challenges of Data Management adoption within DoD?[1]

General Principles

Data management, like other asset management processes, requires understanding available data and leveraging it to achieve organizational goals. It involves balancing strategic and operational needs, which can be achieved by adhering to key principles that guide effective data management practices. A list of 13 general principles, reproduced from the DAMA DMBOK, is provided below that can help shape data management practices across an organization.

1. **Data is an asset with unique properties:** Data is an asset, but it differs from other assets in important ways that influence how it is managed. The most obvious of these properties is that data is not consumed when it is used, as are financial and physical assets.

2. **The value of data can and should be expressed in economic terms:** Calling data an asset implies that it has value. While there are techniques for measuring data's qualitative and quantitative value, there are not yet standards for doing so. Organizations that want to make better decisions about their data should develop consistent ways to quantify that value. They should also measure both the costs of low-quality data and the benefits of high-quality data.

3. **Managing data means managing the quality of data:** Ensuring that data is fit for purpose is a primary goal of data management. To manage quality, organizations must ensure they understand stakeholders' requirements for quality and measure data against these requirements.

4. **It takes Metadata to manage data:** Managing any asset requires having data about that asset (number of employees, accounting codes, etc.). The data used to manage and use data is called Metadata. Because data cannot be held or touched, to understand what it is and how to use it requires definition and knowledge in the form of Metadata. Metadata originates from a range of processes related to data creation, processing, and use, including architecture, modeling, stewardship, governance, Data Quality management, systems development, IT and business operations, and analytics.

5. **It takes planning to manage data:** Even small organizations can have complex technical and business process landscapes. Data is created in many places and is moved between places for use. To coordinate work and keep the end results aligned requires planning from an architectural and process perspective.

6. **Data management is cross-functional;** it requires a range of skills and expertise: A single team cannot manage all of an organization's data. Data management requires both technical and non-technical skills and the ability to collaborate.

[1] Except for where noted, all source material for this chapter comes from the DAMA-DMBOK.

7. **Data management requires an enterprise perspective:** Data management has local applications, but it must be applied across the enterprise to be as effective as possible. This is one reason why data management and data governance are intertwined.

8. **Data management must account for a range of perspectives:** Data is fluid. Data management must constantly evolve to keep up with the ways data is created and used and the data consumers who use it.

9. **Data management is lifecycle management:** Data has a lifecycle and managing data requires managing its lifecycle. Because data begets more data, the data lifecycle itself can be very complex. Data management practices need to account for the data lifecycle.

10. **Different types of data have different lifecycle characteristics:** And for this reason, they have different management requirements. Data management practices have to recognize these differences and be flexible enough to meet different kinds of data lifecycle requirements.

11. **Managing data includes managing the risks associated with data:** In addition to being an asset, data also represents risk to an organization. Data can be lost, stolen, or misused. Organizations must consider the ethical implications of their uses of data. Data-related risks must be managed as part of the data lifecycle.

12. **Data management requirements must drive Information Technology decisions:** Data and data management are deeply intertwined with information technology and information technology management. Managing data requires an approach that ensures technology serves, rather than drives, an organization's strategic data needs.

13. **Effective data management requires leadership commitment:** Data management involves a complex set of processes that, to be effective, require coordination, collaboration, and commitment. Getting there requires not only management skills, but also the vision and purpose that come from committed leadership.

The general principles enumerated above can provide organizations with a strong foundation in their efforts towards impactful data management. But due to the distinct characteristics of data itself, data management has specific challenges that can inhibit the impact and effectiveness of the general principles set forth above. A list of such challenges, organized by the 13 general principles from the DAMA-DMBOK, are described below.

1. **Data differs from other assets:** Data is dynamic, intangible, easily copied, and transportable, yet it is not easy, and sometimes impossible, to reproduce when destroyed or lost and when data is used it is not consumed. These attributes make data management a unique challenge that requires extreme care due to the vital role that data plays in terms of how business is conducted and how an organization functions.

2. **Data valuation:** The benefits and costs of data often lack standardization and consistency across organizations making calculating the value of data complicated. Additionally, data is contextual and

temporal creating further challenges to valuation. It is critical that organizations establish methods to connect financial value with data to ensure informed and consistent decision making.

3. **Data quality:** Data must be reliable. Otherwise, the time and effort to collect, store, secure, and utilize data is pointless. Poor quality data can generate negative consequences for organizations and is additionally a cost, sometimes quite significant. High quality data requires commitment and planning to ensure quality is a part of systems and processes.

4. **Planning for better data:** Planning is necessary to derive value from data and it begins with an understanding that how data is obtained and created is within an organization's control. Short term and long-term goals must be balanced as well as organizational pressures, in addition to time and money, to ensure that they do not hinder adequate planning.

5. **Metadata and data management:** Metadata is a type of data and therefore must also be managed. Poor data management is often associated with a lack of Metadata management at all. A focus on Metadata management can provide a basis for making improvements in overall data management.

6. **Data management is cross-functional:** Different parts of an organization are responsible for managing data at different phases of the data lifecycle. Wide ranging skills and viewpoints are essential to effective data management. An appreciation and understanding of how each piece can work in tandem is vital but integrating these diverse pieces of the larger whole can be a challenge.

7. **Establishing an enterprise perspective:** Data is often disparate and sometimes unique with different parts of an organization. Knowledge and understanding of the breadth and depth of data across the organization is paramount. A lack of such knowledge and understanding inhibits effective use of data enterprise wide.

8. **Accounting for other perspectives:** The creators and collectors of data are often not the eventual users of the data. Understanding of the potential uses of data can help ensure higher quality data and better enable effective planning for the data lifecycle. Additionally, the misuse of data is a risk that should be accounted for to limit the possibility of improper use of data.

9. **The data lifecycle:** Data has a lifecycle and is often dynamic. It also has lineage (sometimes referred to as the data chain). Lifecycle and lineage are often complex within an organization. Focusing on the most critical points in the lifecycle (creation and usage) as well as the most critical data can help organizations navigate some of the complications that can arise over the lifecycle of data.

10. **Different types of data:** Different types of data have different requirements for lifecycle management that can make data management inherently more complicated. Classification and control of different types of data can help mitigate some of the complexity that arises from the different roles, risks, and requirements of different types of data.

11. **Data and risk:** Data represents risk as it has the potential to be misused and misunderstood. One of the biggest risks is information gaps that can create enterprise liabilities resulting in significant hazards to operational efficiency and effectiveness.

12. **Data management and technology:** Business and technical skills are necessary to successfully navigate the wide range of activities that constitute data management. Technology and data are intertwined and necessary to each other as opposed to being in tension with one another. Understanding the impact of technology on data is important such that data needs and business strategy drive technological decisions instead of technology driving decisions about data.

13. **Effective data management requires leadership and commitment:** Data management is not simple or easy and often requires cultural change to be done well. Organizations often underestimate the level of effort required for good data management; lack a strategic approach to data; mismanage and confuse data and information technology; and are not aware of what data is most critical or available. Involvement across the organization coupled with committed leadership is critical to navigating these challenges.

Data Governance

Data governance "provides direction and oversight for data management by establishing a system of decision rights over data that accounts for the needs of the enterprise."

Enabling data to be managed as an asset by an organization is the goal of data governance. To achieve this goal, data governance should be sustainable, embedded, and measured. All three of these qualities of data governance are likely to be moderately difficult to implement within DoD. Sustainability will necessitate organizational commitment and be an ongoing process. It does not require major upheaval, but it does require a measured approach to ensure change and will be highly dependent upon sponsorship, ownership, and leadership. Existing business activities will require data governance to be incorporated into current practices which may prove difficult, especially for a large bureaucratic organization with many levels such as DoD. Establishing a baseline with the expectation for improvement will require establishing demonstrable measures, especially measures to gauge financial impact.

The principles of data governance include leadership and strategy; business-driven; shared responsibility; multi-layered; framework-based; and principle-based. Creative and dedicated leadership is the starting point for data governance. Data governance must be governed by business decisions that shape the interactions of information technology and data. All data stewards and those involved in data management have a shared responsibility to ensure successful data governance. Enterprise wide, local, and everywhere in between will experience and support data governance. Accountabilities and interactions must be defined as coordination is required due to data governance occurring across functional areas. Formal principles to guide data governance will be essential to help mitigate possible resistance.

All of the data governance principles are likely to have a relatively easy implementation path considering much of what each principle entails is already being occurring at some level presently within DoD. But it should be noted that there is almost certainly going to be a learning curve as well as some level of resistance considering data governance will introduce new techniques for decision making. A cultural shift will be required for long

lasting and effective data governance. Ignoring culture reduces the likelihood of success and therefore, adequate attention should be given to managing change during implementation.

Data Architecture

Data architecture "defines the blueprint for managing data assets by aligning with organizational strategy to establish strategic data requirements and designs to meet these requirements."

Connecting technology execution with business strategy is the goal of data architecture. The DAMA-DMBOK does not explicitly identify principles for data architecture. Therefore, with limited guiding principles to rely upon, organizations should ensure that they hire capable and talented architects.

Data architecture implementation should involve at least two components launched simultaneously or in parallel and can begin within a specific part of the organization or a data domain (e.g., product or customer data). The implementation can expand over time as maturity grows. Early projects require more intensive data architecture work before reusable artifacts are available, potentially benefiting from special funding. Data models and artifacts are typically developed within projects and later standardized by data architects. The implementation strategy is significantly influenced by business drivers and must be adaptable in solution-oriented cultures that embrace disruptive technology. An agile approach allows incremental development, with high-level subject area models guiding more detailed work in agile sprints. However, early engagement of data architects is critical to keep up with rapid innovation.

Data Modeling and Design

Data modeling and design ""is the process of discovering, analyzing, representing, and communicating data requirements in a precise form called the data model."

Documenting and validating various perspectives to create applications aligned with current and future business needs, as well as establishing a solid foundation for large scale initiatives to be completed successfully is the goal of data modeling. Potential benefits include reduced support costs, enhanced reusability for future projects, and lowers expenses associated with developing new applications. Additionally, data models serve as a crucial form of Metadata.

Data modeling and database design standards guide business data management, ensuring alignment with Enterprise and Data Architecture while maintaining data quality. These standards should be collaboratively developed by data architects, analysts, and administrators, complementing existing IT standards. Best practices in naming conventions include publishing naming standards for modeling and database objects, ensuring names are unique and descriptive; utilizing logical names that are meaningful to business users while avoiding unnecessary abbreviations; remaining consistent across environments; and using class words (e.g., Quantity, Name, Code) to distinguish attributes and columns as this will assist in data analysis.

Best practices in database design include performance and ease of use; reusability; integrity; security; and maintainability. Approved users will need quick and easy access in a form that is usable and relevant to the business activity as such actions will maximize business value. Database structure should allow for multiple applications to utilize data while also allowing the data to serve multiple purposes. Data should consistently

maintain business relevance and accuracy across all contexts and any violations should be detected and reported immediately. Accurate data should be readily available to authorized users while considering the privacy concerns of all stakeholders. The value of data to the organization should exceed the cost of data work, including all aspects, while ensuring future changes to business processes and needs can be executed in the most expedient form possible. All of these best practices should be fairly easy to implement at DoD as all of these practices are currently in operation in one form or another across the organization.

Data Storage and Operations

Data storage and operations includes the design, implementation, and support of stored data to maximize its value. Operations provide support throughout the data lifecycle from planning for to disposal of data.

The principles for Data Storage and Operations include using automation opportunities, building for reuse, applying best practices, connecting database standards, and setting expectations for the Database Administrator's role in projects. To use automation opportunities, database development processes should be automated, tools developed, and processes shortened to reduce errors and pressure on the development team. Building with reuse in mind refers to using abstracted and reusable data objects, such as database views, triggers, or functions and stored procedures, that prevent applications from being coupled to database schemas. In applying best practices, database administrators should follow standards and best practices but deviate when necessary. Furthermore, database standards should be connected to support requirements, for example with Service Level Agreements which should reflect the database administrators' and developers' ideas. Finally, the database administrator should understand the needs and expectations of the project in order to better understand their role, even having a second administrator to clarify expectations when necessary.

The most difficult principles to implement would be using automation opportunities and setting expectations for database administrators because of the computer power, manpower, and coordination required. Although automation opportunities are covered extensively in the DoD strategies, they require great computer power to implement, as every step would be run by a computer for all data and records management. Setting expectations for database administrators would also be difficult because it is not included in current strategies, leaving no path to implementation. Additionally, this principle requires a second database administrator to clarify expectations to the primary administrator, which requires unnecessary manpower and collaboration between the team.

The easiest principle to implement would be applying best practices, as it is covered in DoD strategies and is already being implemented. Best practices are referenced under the principle of Data Ethics as a way to promote more ethical data use [Department of Defense, 2020]. Additionally, some best practices are already being implemented, including employing agile development, building intuitive interfaces for human adoption of technology, developing products for customer needs, having product portfolios with shared digital foundations, and experimenting with minimum viable products to improve them [DoD CDAO Strategy, 2023]. The success of this implementation proves that this principle would be the easiest to continue to implement.

The principles of using automation opportunities, building for reuse, and applying best practices align with current DoD data strategies. The principle of using automation opportunities is referenced in the DoD data strategy's principles of Data Collection and Design for Compliance [Department of Defense, 2020]. In Data

Collection, the strategy states that the collection process should be automated as much as possible to reduce human error, while Design for Compliance focuses on automating the information management lifecycle to secure data and records [Department of Defense, 2020]. Additionally, jointly interoperable infrastructure for data, analytics, and AI capabilities development will be automated [DoD CDAO Strategy, 2023]. Building for reuse also overlaps with DoD strategies' goal to Make Data Accessible, as it plans to control data access and sharing through reusable Application Programming Interfaces [Department of Defense, 2020]. Additionally, this principle aligns with the strategy to create a culture of data sharing and reuse to collaborate across domains and destruct data silos [DoD CDAO Strategy, 2023]. Promoting best practices is mentioned briefly in the principle of Data Ethics, where it is states that ethical data use will be supported by identifying and promoting the use of best practices [Department of Defense, 2020]. Additionally, best practices from technology developers are used to employ agile development, build intuitive interfaces, develop products for customer needs, and more [DoD CDAO Strategy, 2023].

Data Security

Data security ensures that data privacy and confidentiality are maintained, that data is not breached, and that data is accessed appropriately.

Principles of Data Security include collaboration, enterprise approach, proactive management, clear accountability, metadata-driven approach, and risk reduction by reducing exposure. For collaboration, IT security administrators, data stewards, the data governance community, internal and external auditors, and the legal department should work together for data security. The enterprise approach refers to standards and policies being applied organization-wide. Proactive management requires engaging stakeholders, managing change, and persevering through organizational or cultural problems like organization of responsibilities. This can be remedied with clear accountability, which defines roles and responsibilities of data management. Security is also aided by being metadata-driven, or having security classification for data elements, and reducing exposure of sensitive data.

The principles that would be the most difficult to implement are proactive management, collaboration, and risk reduction by reducing exposure. Proactive management would be difficult because it requires defining the roles and responsibilities of many stakeholders, including in information security, information technology, data administration, and business, who need to collaborate for management. Additionally, there are no plans of how to achieve this in current DoD data strategies. Collaboration would also be difficult to implement because it requires security administrators, data stewards, governance, auditors, and legal professionals to cooperate and work together. However, there are guides for how this could be facilitated, like through the CDO Council [Department of Defense, 2020], in the data strategies. Furthermore, risk reduction by reducing exposure is a difficult task because there are no guidelines in current strategies to steer this effort.

The easiest principle to implement would be a metadata-based approach because it simply requires attaching security classifications to data elements. Although this would require manpower, if it is performed in data catalogs as in the Make Data Understandable goal [Department of Defense, 2020], it would be a simple task to complete.

The principles of collaboration, enterprise approach, clear accountability, and a metadata-driven approach follow the guide of the DoD data strategies. For collaboration, the strategies identify promoting Talent and Culture through supporting collaboration among data experts as an essential capability [Department of Defense, 2020]. This will be supported with the CDO Council's role as a space for data officers from various domains to collaborate [Department of Defense, 2020]. Additionally, the strategy emphasizes collaboration with industry, for example to develop and deploy AI-enabled systems for use in Joint Warfighting [DoD CDAO Strategy, 2023]. Furthermore, the principle of enterprise approach is aligned with strategies to strengthen data governance. This is supported by the CDO Council's role in overseeing policy and data standards organization-wide [Department of Defense, 2020]. The principle of clear accountability aligns with the guiding principle of Collective Data Stewardship, which states that the DoD is defining roles and responsibilities for data stewardship such as data stewards, custodians, and managers [Department of Defense, 2020]. In addition, the principle of a metadata-driven approach follows the goal of to Make Data Understandable, which requires data to be inventoried in catalogs with information on security, limitations, and restrictions on use [Department of Defense, 2020].

Data Integration and Interoperability

Data integration and interoperability includes processes related to the movement and consolidation of data within and between data stores, applications, and organizations.

Principles relevant to Data Integration and Interoperability are taking an enterprise perspective, balancing local and enterprise data needs, and ensuring business accountability. The principle of taking an enterprise perspective means to implement this perspective incrementally for future extensibility. For balancing local and enterprise data needs, this should be performed throughout support and maintenance. In addition, business experts should be involved in designing or changing data transformation rules in order to facilitate business accountability.

The principles that would be the most difficult to implement are taking an enterprise perspective and ensuring business accountability. For taking an enterprise perspective, this principle does not align with data strategies and it is therefore unclear how this incremental implementation should occur. Furthermore, the principle of ensuring business accountability is also not mentioned in strategies, and requires expertise from business professionals, as well as cooperation with those managing data transformation rules.

The easiest principle to implement would be balancing local and enterprise data needs as it is represented in data strategies. The data strategies provide a clear example of how this balance could occur in the example of data quality.

The principle of balancing local and enterprise data needs aligns with current DoD data strategies. For example, the strategy states that local requirements are the foundation of data quality [DoD CDAO Strategy, 2023]. In addition, the strategy explains that in its goal to increase data quality and access for advanced analytics and AI capabilities, it will focus on collecting, storing, and managing data relevant to enterprise needs [DoD CDAO Strategy, 2023]. This example in data quality demonstrates the DoD's effort to balance local and enterprise data needs.

Document and Content Management

Document and content management includes planning, implementation, and control activities used to manage the lifecycle of data and information found in a range of unstructured media, especially documents needed to support legal and regulatory compliance requirements.

The principles of Document and Content Management are organization-wide accountability and engaging experts in policy and planning. Additionally, there are several relevant principles in ARMA International's Generally Acceptable Recordkeeping Principles, including: accountability, integrity, protection, compliance, availability, retention, disposition, and transparency. The first principle refers to assigning everyone in an organization a role in creating, using, retrieving, and disposing of records in compliance with proper procedures. The next principle suggests involving records management professionals, or other employees like business stewards, in policy and plans for industry and legal jurisdiction, in order to form an effective approach to records management. For the ARMA principles, accountability suggests assigning a senior executive to guide staff in executing policies and processes, ensuring program auditability. In addition, integrity refers to creating an information governance program to ensure authenticity and reliability of information. This information governance program should also protect personal information and comply with laws or organizational policies. For availability, information should be easily accessible, even past information which should be retained for an appropriate time based on operational, legal, or fiscal requirements. Furthermore, organizations should provide secure disposition of information as necessary, and should document policies and processes to staff or other stakeholders to maintain transparency.

The most difficult principle to implement would be engaging experts in policy and planning. This principle is not mentioned in data strategies, leaving no realistic path to implementation. Additionally, this requires the expertise of records management professionals, or the manpower to train others to become data stewards. This also requires coordination between those creating the policies and plans and the experts to create equitable rules.

The easiest principles to implement would be organization-wide accountability and accountability, integrity, and compliance. The principle of organization-wide accountability is heavily covered in data strategies under Collective Data Stewardship and Expanding Digital Talent Management for data quality, but there are even more examples in other domains [Department of Defense, 2020]. Additionally, the ARMA principles of accountability, integrity, and compliance of the information governance program would be easy to implement because something similar is happening with the DoD CDO, CIO, and CDO Council [Department of Defense, 2020]. These bodies manage the same aspects of data management as the proposed information governance program, providing a real example of how this is and will continue to be implemented.

The principles of organization-wide accountability, general accountability, integrity, compliance, availability, retention, and disposition align with current DoD data strategies. For organization-wide accountability, data strategies refer to this principle in the guiding principle of Collective Data Stewardship [Department of Defense, 2020] and the goal to Expand Digital Talent Management [DoD CDAO Strategy, 2023]. Collective Data Stewardship defines roles and responsibilities, separating members into data stewards, custodians, and managers [Department of Defense, 2020]. The goal to Expand Digital Talent Management builds on this by

predicting increases in roles such as data architects, stewards, and user experience designers [DoD CDAO Strategy, 2023]. The ARMA's principle of accountability aligns with the DoD strategy to strengthen governance by increasing oversight from the CDO, CIO, and the CDO Council [Department of Defense, 2020]. The CIO will ensure compliance with CDO guidance on IT investments, DoD policy, and budgeting, while the CDO Council will oversee policy and data standards of the DoD [Department of Defense, 2020]. These governing bodies also may promote the principle of integrity by overseeing data management efforts, and the principle of compliance by complying with the CDO's guidance [Department of Defense, 2020]. The principle of availability is also relevant in DoD data strategies through the guiding principle of Enterprise-Wide Data Access and Availability, which emphasizes all authorized individuals having access to appropriate data [Department of Defense, 2020]. Furthermore, the principle of retention is mentioned in an objective of the goal to Make Data Secure, where it requires the development and implementation of content and record retention rules [Department of Defense, 2020]. The principle of disposition also aligns with this goal, where it requires granular privilege management to be implemented to govern the disposition of data, as well as disposition to be audited [Department of Defense, 2020].

Reference and Master Data

Reference and master data include ongoing reconciliation and maintenance of core critical shared data to enable consistent use across systems of the most accurate, timely, and relevant version of truth about essential business entities.

The principles for Reference and Master Data management are shared data, ownership, quality, stewardship, controlled change, and authority. These principles state that data should be shareable organization-wide, and belong to the organization. Additionally, there should be regular data quality monitoring, with Business Data Stewards ensuring the Reference Data's quality. Furthermore, changes to Master Data must be reversible and conducted with oversight, which is similar to the requirement for Reference Data to only have approved changes. Finally, Master Data should only be replicated from the system of record.

The principles of stewardship and controlled change would be the most difficult to implement because they require a high level of expertise, manpower, coordination, and in the case of controlled change, are not aligned with current DoD data strategies. Although stewardship is aligned with the DoD's principle of Collective Data Stewardship, implementing this principle requires several different specialized employees, including stewards, custodians, and managers [Department of Defense, 2020]. This requires manpower and expertise, but also coordination between the different positions who decide policies, enforce them, and manage data quality [Department of Defense, 2020]. On the other hand, controlled change does not align with current DoD data strategy, and also requires manpower and computing power. This is because controlled change relies on all changes being reversible, requiring computer power, and oversight on changes, requiring manpower for each small change to Master Data.

Although the principles of shared data and data quality require coordination and manpower, these are the easiest to implement because of their overlap with current DoD strategies. The principle of shared data is aligned with the principle of Enterprise-Wide Data Access, and data quality to the objective of Data Quality

[Department of Defense, 2020]. The principle and objective from the data strategy covers the goal of what to do, as well as how to do it, facilitating the easiest implementation.

DoD data strategies address the principles of shared data, quality, and stewardship. The DoD references the principle of shared data in their principle of Enterprise-Wide Data Access and Availability [Department of Defense, 2020], where it emphasizes the responsibility to provide data and to share information broadly [Department of Defense, 2020]. Furthermore, data sharing is facilitated by treating data as a product [DoD CDAO Strategy, 2023]. The strategies also cite Data Quality as an objective achieved by using data quality management techniques [Department of Defense, 2020], such as implementing a decentralized network among data providers and users [DoD CDAO Strategy, 2023]. In addition, the DoD shares a guiding principle of Collective Data Stewardship with the DMBOK, where it requires assigned data stewards, custodians, and managers to establish policies, promote data value and enforce policies, and manage daily quality, respectively [Department of Defense, 2020].

Data Warehousing and Business Intelligence

Data warehousing and business intelligence includes the planning, implementation, and control processes to manage decision support data and to enable knowledge workers to get value from data via analysis and reporting.

Principles for Data Warehousing and Business Intelligence include focusing on business goals, starting with the end in mind, designing globally but building locally, summarizing last, transparency, building metadata in a warehouse, collaborating, and using the right tools for different data consumers. For business goals, the data warehouse should respond to organizational priorities and problems, as well as the business priority to start with the end in mind. Additionally, for a global design and local build, building and delivery should occur incrementally through focused sprints for faster return on investment. Summary should occur last and not be used to fill in missing details. Metadata should be built with the warehouse to give consumers better value and information, as well as to be transparent and inform stakeholders. Furthermore, collaboration is encouraged with other initiatives in Data Governance, Quality, and Metadata, yet it is imperative to use the specific tools relevant to each group of data consumers.

The principles that would be most difficult to implement are designing globally but building locally and collaboration. This is because designing globally but building locally is misaligned with current DoD data strategies, and requires the expertise and power for focused sprints. In addition, collaboration requires coordination of multiple communities, including the Data Governance, Quality, and Metadata groups in the DAMA-DMBOK, and between the data governance and operational communities for Joint All Domain Operations [Department of Defense, 2020]. However, collaboration is aligned with DoD data strategies, making it partially easier to implement.

The easiest principles to implement would be focusing on business goals and building metadata in a warehouse, because of their alignment with DoD data strategies and the lack of coordination necessary. For focusing on business goals, the DoD CIO is the sole authority tasked with integrating data priorities into the Digital Modernization program to sync with other efforts [Department of Defense, 2020]. This requires less manpower and coordination than other principles. Building metadata in a warehouse requires some computer and possibly

manpower to complete all the different metadata tasks [Department of Defense, 2020], but the plan for implementation is well laid out in the strategy, making it easy to implement.

DoD strategies align with some of these principles including focusing on business goals, starting with the end in mind, building metadata in a warehouse, collaboration, and using the right tools for different data consumers. To focus on business goals, the DoD CIO will ensure that data priorities are integrated into the DoD Digital Modernization program to sync with the cloud, AI, Command, Control, Communications, and cybersecurity efforts [Department of Defense, 2020]. Furthermore, the DoD will challenge vendors to solve specific business and mission problems, starting with the end in mind [DoD CDAO Strategy, 2023]. In addition, the DoD has an objective to supply data with protection, lineage, and pedigree metadata to make data valuable to users and stakeholders, which is related to the principle of building metadata in a warehouse [Department of Defense, 2020]. The strategy also references collaboration under the focus area of Joint All Domain Operations, where collaboration must occur between the data governance and operational communities [Department of Defense, 2020]. Furthermore, the strategy lists seven goals in making data best fit to different data consumers, including: visibility, accessibility, understanding, links, trustworthiness, interoperability, and security (VAULTIS) [Department of Defense, 2020].

Metadata

Metadata includes planning, implementation, and control activities to enable access to high quality, integrated Metadata, including definitions, models, data flows, and other information critical to understanding data and the systems through which it is created, maintained, and accessed.

Metadata management follows principles of organizational commitment, strategy, enterprise perspective, socialization, access, quality, audit, and improvement. This means that metadata management requires senior management support and funding. Additionally, metadata management needs a strategy for how metadata will be created, maintained, integrated, and accessed, in alignment with business priorities and to drive requirements. Furthermore, an enterprise perspective is important for future extensibility, but it is necessary to implement it in incremental delivery. Communication of the necessity and purpose of the metadata, availability to staff, quality assurance from production owners, audits of metadata standards, and a feedback system for consumers are also important.

The principles of organizational commitment, enterprise perspective, and socialization would be the most difficult to implement because of their misalignment with DoD data strategies and need for coordination. These principles are not referred to in data strategies with plans for implementation, which is especially problematic for the principle of enterprise perspectives because of its vague objective of future extensibility. Additionally, organizational commitment requires the support of senior management and funding, which relies heavily on coordination between management and employees involved in metadata management to get support for their needs. Furthermore, socialization requires coordination to communicate the need and purpose of metadata from creators to users.

The easiest principles to implement would be strategy, access, and improvement because of their alignment with DoD strategies and clear paths to implementation. These principles mirror the DoD strategies' objective to create a metadata strategy for implementing standards [Department of Defense, 2020], principle of Enterprise-

Wide Data Access and Availability [Department of Defense, 2020], and priority of feedback from consumers [DoD CDAO Strategy, 2023].

The DoD shares principles in strategy, access, and improvement in its data strategies. For strategy, the DoD has objectives to implement metadata standards like location and access methods, common syntax, and how to join and integrate data to standardize access, creation and maintenance, and integration [Department of Defense, 2020]. In addition, the DoD follows a principle of Enterprise-Wide Data Access and Availability, which is similar to the DMBOK's principle of making data available to staff [Department of Defense, 2020]. Finally, the DoD also prioritizes feedback from data consumers, especially through rapid feedback loops to meet user demands [DoD CDAO Strategy, 2023], in alignment with the principle of data improvement [Department of Defense, 2020].

Data Quality

Data quality includes the planning and implementation of quality management techniques to measure, assess, and improve the fitness of data for use within an organization.

Relevant principles to data quality include criticality, lifecycle management, prevention, root cause remediation, governance, a standards-driven approach, objective measurement and transparency, embedding in business processes, systematic enforcement, and connection to service levels. For criticality, data quality should focus on the highest risk, most critical enterprise data. In addition, data should also be managed across the entire lifecycle and between systems. Programs should also prevent data errors, addressing the root causes in processes or systems. Furthermore, high quality data development should be supported by data governance, and vice versa. Stakeholders should identify measurable requirements for data, with data quality levels being measured objectively and consistently. Business owners should be responsible for the quality of data output from their processes, and system owners need to enforce requirements. Finally, the management of reporting and issues should also be incorporated into Service Level Agreements (SLA).

The most difficult principles to implement would be governance and objective measurement and transparency because of the coordination and manpower required. Although the principle of governance aligns with the DoD data strategies' goal to focus on data governance and quality at the same time [DoD CDAO Strategy, 2023], this is difficult to do because it requires coordination of personnel on both sides to work in sync. In addition, objective measurement and transparency relies on manpower to perform continuous measurements and check against each other for objectivity. This principle also is misaligned with DoD data strategies' objectives.

The principles of lifecycle management and connection to service levels would be the easiest to implement because they overlap with DoD data strategies and have clear paths to implementation. Lifecycle management is referred to in the principle of Data for Artificial Intelligence Training [Department of Defense, 2020], Design for Compliance [Department of Defense, 2020], and has a path to implementation through assessment using data quality measures and VAULTIS [DoD CDAO Strategy, 2023]. Additionally, connection to service levels is connected to the DoD principle of solving problems at the lowest level [Department of Defense, 2020].

The DoD follows the principles of criticality, lifecycle management, governance, and connection to service levels as listed in their data strategies. For criticality, the DoD explains in the problem statement that quality, critical data is not accessible, so efforts like data and records management for critical data are important to improve

this [Department of Defense, 2020]. In addition, the DoD has a principle in Collective Data Stewardship that aligns with the principle of lifecycle management [Department of Defense, 2020]. In this, the data stewards, custodians, and managers have to achieve accountability throughout the lifecycle [Department of Defense, 2020]. Lifecycle management is also important in the principle of Data for Artificial Intelligence Training, where datasets for AI training and algorithmic models must be managed across the lifecycle [Department of Defense, 2020]. Additionally, lifecycle management is relevant to the Design for Compliance principle, where IT solutions should automate the information management lifecycle [Department of Defense, 2020]. Furthermore, this management of data quality across the lifecycle may include assessment using data quality dimensions and the VAULTIS framework [DoD CDAO Strategy, 2023]. The strategies also state that data governance and data quality should be focused on at the same time to mitigate risks, including the replication of unintended bias enterprise-wide, which aligns with the governance principle [DoD CDAO Strategy, 2023]. The DoD also abides by the principle of connection to service levels by resolving issues at the lowest level possible [Department of Defense, 2020].

Big Data

The only formed principle related to Big Data management is careful metadata management for accurate information on files, origins, and value.

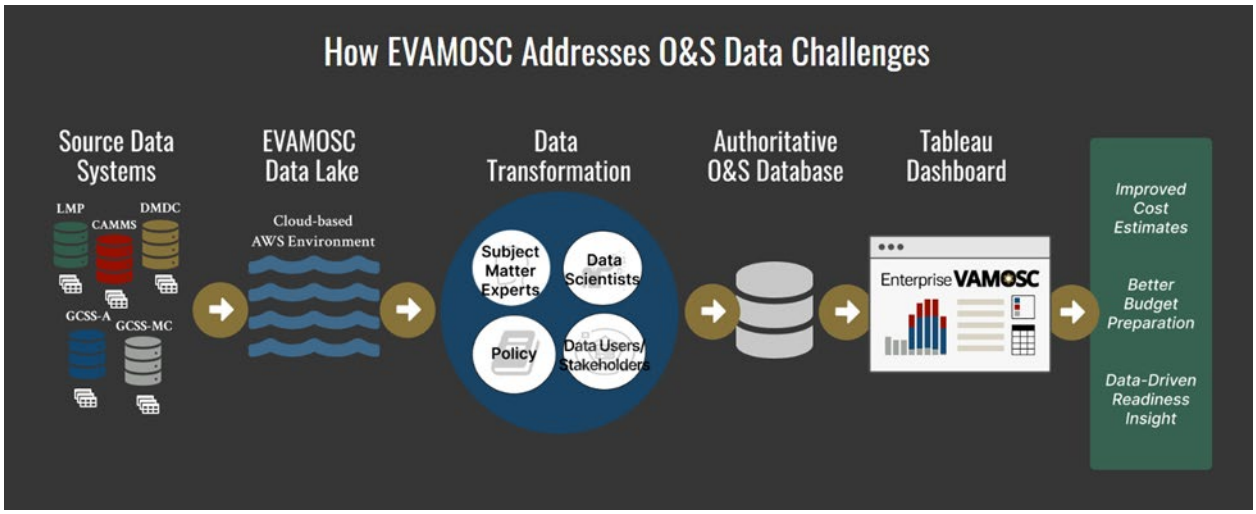
This principle would be easy to implement because it has clear links to DoD objectives and has a plan for implementation in metadata standards [Department of Defense, 2020]. These standards would achieve the principle's goal, yet may require man or computer power to maintain consistently accurate information.

Big Data's principle mirrors the DoD's objectives for metadata management. For example, the data strategy contains objectives to implement metadata standards including location and access methods, common syntax, and integration processes, which help users access and find accurate information on files [Department of Defense, 2020].

7. Recommended pathways to implementation of relevant commercial data management best practices, standards, and solutions

DoD's IP Cadre should consider using the existing EVAMOSC model as a starting place for their own data management efforts as such an approach is likely to have the easiest path to implementation. EVAMOSC utilizes a five-layer process composed of sourcing data systems, establishing a data lake, data transformation, building an authoritative database, and a dashboard. Figure 1 provides a visualization of this process and the components. Considering the unique challenges that arise when managing cost data, the complexity and security requirements of EVAMOSC's model may be fairly analogous to those related to managing IP data and make such a model potentially ideal as a foundational framework.

Figure 1. EVAMOSC Model



The first component of the EVAMOSC model is sourcing data systems. DoD requires secure, scalable, and compliant data systems. More details about the benefits. Therefore, DoD will need to:

- Define clear data system requirements because a well-defined data strategy prevents inefficiencies, data silos, and compliance risks. Best practices include identifying key use cases, determining data types, planning for data interoperability, and assessing scalability needs.
- Prioritize security, compliance, and access controls because IP data includes sensitive, classified, and proprietary assets that must be protected. Best practices include, following NIST 800-53 and FedRAMP guidelines, implementing role-based access control (RBAC), using encrypted storage and backups and applying blockchain for IP tracking.
- Select the right database and cloud infrastructure (part of component #2) because DoD needs reliable, scalable, and compliant data storage solutions. Best practices include using federally approved cloud services, selecting a suitable database type, enabling hybrid and on-premises integration, optimizing storage for performance and cost.

- Ensure data interoperability and integration capabilities because DoD must share IP and licensing data across multiple stakeholders. Best practices include implementing open APIs and standards, ensuring compatibility with external systems, adopting common data models, and enabling secure data federation.

- Implement automated data processing and AI/ML analytics because automation reduces manual data entry, errors, and compliance risks. Best practices include using ETL pipelines for data ingestion, leveraging AI for metadata extraction, enabling real-time monitoring and alerts, and applying machine learning for IP trend analysis.

- Ensure strong data governance and auditing because a structured governance framework ensures data accuracy, accountability, and compliance. Best practices include defining clear data stewardship roles, enabling automated audit logging, standardizing data cataloging and tagging, establishing data retention and disposal policies.

- Optimize user experience and self-service access because a well-designed system improves adoption, usability, and efficiency. Best practices include implementing searchable dashboards, supporting custom reports and visualizations, enabling self-service analytics, and ensuring mobile and web accessibility.

- Plan for scalability and future expansion because as DoD acquires more IP data, systems must scale effectively. Best practices include using cloud-native architectures, implementing containerization, planning for AI and/or blockchain integration, and ensuring vendor independence.

The second component of the EVAMOSC model is establishing a data lake. A well-implemented data lake serves as a foundation for enterprise-wide data management and analytics. The key to success lies in choosing the right architecture, enforcing governance, optimizing costs, and leveraging advanced analytics to drive business insights. DoD requires secure, scalable, and compliant data lakes to store, process, and analyze large volumes of structured and unstructured data. Therefore, DoD will need to:

- Compare data lakes by evaluating security and compliance, data governance, data formats and interoperability, scalability and performance, cost and vendor lock in, and integration and compatibility.

- Select the right data lake architecture for IP management considering the needs for large-scale data processing, compliance tracking, and analytics. Best practices include ensuring proper cataloging of licensing agreements, expirations dates, and IP rights; establishing tiered storage (e.g., cold storage for expired licenses, active storage for current agreements); and accommodating dynamic and evolving license agreement formats particularly SNLRs; enabling batch (ETL) and real-time (streaming) ingestion from multiple sources (legal databases, financial systems, licensing records).

- Mitigate common challenges by implementing RBAC and ABAC, using automated data validation and cleansing, enforcing encryption, audit logging, and access monitoring and using object storage, indexing, and partitioning for optimization. Best practices include encrypting sensitive data at rest and in transit (AES-256, TLS 1.2+); implementing role-based access controls (RBAC) for IP and license stakeholders; maintaining immutable audit logs for all license modifications, renewals, and sublicensing activities; and using AI-powered compliance monitoring to detect unauthorized sublicensing or expired agreements.

- Use automated ETL pipelines to continuously update IP license records in real-time. Data ingestion strategies include batch processing (ETL), real-time streaming, and a hybrid approach. Recommended

tools for data processing include Apache NiFi/AWS Glue, Apache Spark/Databricks, and PostgreSQL/Snowflake.

- Leverage AI, analytics and automation to enhance decision-making in government IP management by automating contract risk analysis and compliance tracking. Best practices include using NLP to extract key clauses from licensing contracts; flagging potential unauthorized sublicensing or fraudulent IP usage; using predictive analytics for forecasting licensing demand & renewal trends.

- Optimize cost and performance. Best practices include implementing tiered storage; enabling auto-scaling for workloads, and using data compression and deduplication.

- Ensure future-proofing and scalability. Best practices include supporting multi-agency collaboration; adopting serverless and cloud-native solutions; enabling multi-cloud and hybrid deployments.

The third component of the EVAMOSC model is data transformation. Effective data transformation requires a blend of technology, expertise, and governance. By engaging SMEs, data scientists, and stakeholders, organizations can ensure data is relevant, accurate, and actionable. Establishing strong policies, automating workflows, and continuously optimizing processes further enhance the value of transformed data. Transforming data involves converting raw data into meaningful, structured, and actionable insights. This process requires collaboration among stakeholders, including subject matter experts (SMEs), data scientists, business users, and IT professionals, while ensuring compliance with data policies and governance frameworks. Therefore, DoD will need to:

- Utilize general best practices such as converting raw legal and licensing data into standardized formats (structured databases, JSON/XML for interoperability); removing duplicate, outdated, or inconsistent licensing data to ensure accuracy; using controlled vocabularies and metadata tagging for easy search and retrieval; and using AI for contract analysis, anomaly detection, and compliance monitoring.

- Collaborate and engage with stakeholders. Best practices include defining clear data ownership; ensuring cross-agency collaboration; providing regular training; and adopting agile data governance models.

- Develop data transformation policies to guide data transformation practices ensuring transparency, compliance, and efficiency. Best practices include defining structured formats for IP records, metadata, and licensing agreements; implementing automated quality checks and require SMEs to validate transformed data; enforcing RBAC and encryption for sensitive IP data; maintaining immutable logs for licensing modifications and approvals; and defining storage duration, archival policies, and automatic deletion of expired records

- Standardize structured and unstructured data. Best practices include converting licensing agreements into structured formats; normalize data across agencies; ensure consistent naming conventions across government IP records; use open data standards; use NLP & AI for legal document processing; metadata tagging for fast retrieval; automate contract arising such as expiration dates, sublicensing terms, and royalty agreements.

- Ensure accuracy, consistency, and completeness through cleansing and quality assurance. Best practices include deduplicating entries using fuzzy matching & AI-based record linkage; using data imputation techniques or request manual SME review; standardizing formats using automated transformation rules

(e.g., standard contract templates); and implementing AI-powered metadata validation to detect mismatches.

- Leverage AI, analytics, and automation. Best practices include extracting contract clauses, IP classifications, and compliance rules; identifying high-risk licenses, unauthorized sublicensing, or compliance violations; digitizing physical contracts and historical IP agreements; forecasting licensing trends, renewal likelihood, and financial impact.

- Implement data transformation pipelines by deploying ETL (Extract, Transform, Load). Best practices include automated data ingestion; enabling real-time and batch processing; applying data encryption and access controls; and ensuring API-driven interoperability.

The fourth component of the EVAMOSC model is the development of an authoritative IP database. A well-designed and managed structured database improves data integrity, performance, and security. By implementing strong schema design, optimizing queries, ensuring security, and leveraging automation, DoD can maximize the efficiency and value of their data. Structured databases, such as relational databases (e.g., MySQL, PostgreSQL, SQL Server, Oracle), require careful planning and management to ensure efficiency, scalability, and security. Therefore, DoD will need to:

- Plan for scalability. Best practices include using a relational model (RDBMS) with properly normalized tables to track IP, licenses, stakeholders, and compliance; implementing partitioning for large datasets to improve query performance; utilizing federated access if integrating with other government databases (e.g., USPTO, DoD, DOE, NSF).

- Define a clean schema design and data structuring strategy. Best practices include using 3NF (Third Normal Form) to eliminate redundancy while ensuring relational integrity; creating indexes on frequently searched fields (e.g., IP Title, License ID, Agency Name); and implementing UUIDs for unique and secure record identification.

- Ensure security and compliance. Best practices include using Role-Based Access Control (RBAC) with defined permissions (e.g., Read-Only for Public, Modify for Admins); implementing Multi-Factor Authentication (MFA) for all government personnel accessing the database; using Single Sign-On (SSO) integration with government identity management systems (e.g., CAC/PIV authentication); and using AES-256 encryption for sensitive data at rest and TLS 1.2+ encryption for data in transit.

- Optimize query performance. Best practices include using materialized views for frequently accessed data (e.g., active government licenses); implementing query caching (e.g., PostgreSQL pg_bouncer) to reduce redundant computations; and leveraging full-text search for efficient lookup of IP titles and legal agreements.

- Consider retention and auditing needs. Best practices include implementing automated archival policies; enabling audit logging to track modifications and access history; and automating license renewals and expiration alerts using scheduled jobs.

- Catalog data and metadata. Best practices include utilizing metadata tagging for easy search and classification of IP rights; and implementing data lineage tracking to ensure proper governance.

- Enable business intelligence and reporting. Best practices include creating real-time dashboards; generating monthly compliance reports for license agreements; and using AI-powered analytics for predicting renewal trends and enforcement risks.

- Plan for backup and disaster recovery. Best practices include setting up multi-region replication (e.g., PostgreSQL Logical Replication, AWS RDS Multi-AZ); automating backups (daily full, hourly incremental) stored in an encrypted environment; and implement failover clustering to ensure continuous availability.

The fifth component of the EVAMOSC model is the development of dashboards and/or applications. Dashboards and data applications serve as critical tools for transforming raw data into actionable insights. By focusing on user needs, performance, security, and analytics, organizations can maximize the value of their data applications. Effective dashboards require careful selection, thoughtful design, and structured management. They enable users to retrieve, analyze, and visualize data efficiently, supporting better decision-making processes. Therefore, DoD will need to:

- Select the right dashboard and application tools. Best practices include identifying key decision-makers and their specific data needs; ensuring the dashboard handles large datasets efficiently; choosing tools that integrate with existing infrastructure (e.g., Power BI for Microsoft environments, Looker for Google Cloud, Tableau for hybrid analytics); and choosing the appropriate deployment model based upon security concerns and business needs.

- Create effective dashboards and applications. Best practices include following the 5-Second Rule; organizing data to convey a clear narrative (e.g., trend analysis before forecasting); automating data validation to prevent reporting errors; pre-aggregating large datasets or using caching to improve speed; supporting executive summaries and detailed drill-downs; and using color-blind-friendly palettes and allow screen reader compatibility.

- Maintain and update dashboards and data applications. Best practices include using ETL tools (e.g., dbt, Apache Airflow) to refresh data without manual intervention; tracking who is using the dashboard and which insights matter most; ensuring dashboards align with data security laws; using materialized views or indexed queries to speed up load times; and gathering input from users for iterative improvements.

- Leverage dashboards for decision-making and analytics. Best practices include enabling self-service analytics to allow non-technical users to generate insights without SQL expertise; applying AI for predictive analytics (e.g., churn prediction, anomaly detection); automating actions based on insights (e.g., trigger alerts when KPIs drop below threshold); notifying stakeholders when key metrics deviate from expected ranges; and training teams on best practices for interpreting and acting on insights.

Conclusion

Challenges to adoption include the complexities of data valuation, enterprise-wide governance, and security concerns. To overcome these obstacles, the report recommends leveraging the Enterprise Visibility and Management of Operating and Support Cost (EVAMOSC) model, establishing secure data lakes, implementing AI-powered data transformation, and developing intuitive dashboards for improved decision-making. Moving forward, the DoD should focus on identifying data sources, evaluating storage solutions, defining security protocols, and formulating an IP-specific data management policy. By integrating commercial methodologies with AI-driven analytics, the DoD can achieve a more secure, interoperable, and efficient data management framework.

By leveraging commercial standards and AI-driven data management techniques, DoD can enhance its IP data governance, security, and interoperability while addressing challenges in implementation and adoption.

Next Steps

1. Research and identify data Sources and data sourcing characteristics.
2. Research and identify solutions and trade space of data storage options (data lake and structured IP database).
3. Research and define data characterization, marking, handling, and security requirements.
    a. What data can be shared, and with whom?
4. Research needs for IP-specific Data Management policy with comparison of existing Data Management strategy/policies and commercial best practices.

References

Adobe Experience Cloud Team. (2023). What is digital rights management (DRM)? Adobe.com. Retrieved December 12, 2024, from https://business.adobe.com/blog/basics/digital-rights-management

Abie, H. (2007, January). Frontiers of DRM knowledge and technology. International Journal of Computer Science and Network Security, 7(1), 216–231. Retrieved December 3, 2024, from http://paper.ijcsns.org/07_book/200701/200701B05.pdf

AltexSoft. (2021, December 20). Data labeling in machine learning: Process, types, and best practices. Retrieved January 16, 2025, from https://www.altexsoft.com/blog/data-labeling/

Amplitude. (n.d.). What is data taxonomy? Retrieved January 14, 2025, from https://amplitude.com/explore/data/what-data-taxonomy#data-taxonomy-full-definition

Anton, P., et al. (2019). Assessing Department of Defense use of data analytics and enabling data management to improve acquisition outcomes. RAND Corporation. Retrieved January 11, 2025, from https://www.acq.osd.mil/asda/dpc/api/docs/anton%20mckernan%20et%20al.%202019%20-%20dod%20data%20analytics%20-%20rand%20rr-3136-osd.pdf

Bailey, C. (2022, December 20). Exploring DRM as an industry standard security addition to file transfers. Cyber Protection Magazine. Retrieved December 13, 2024, from https://cyberprotection-magazine.com/exploring-drm-as-an-industry-standard-security-addition-to-file-transfers.

Brown, J. (2024, February 13). How the DoD unified data organization-wide with Apigee. Google Cloud Blog. Retrieved January 1, 2025, from https://cloud.google.com/blog/topics/public-sector/how-the-dod-unified-data-organization-wide-with-apigee

CCS Learning Academy. (2024, May 21). Top 10 data cleaning techniques and best practices for 2024. Retrieved January 16, 2025, from https://www.ccslearningacademy.com/top-data-cleaning-techniques/

Cohen, J. (2003). The challenge of Digital Rights Management technologies. NIH.gov. Retrieved December 10, 2024, from https://www.ncbi.nlm.nih.gov/books/NBK221850/

DAMA International. (2017). DAMA-DMBOK (2nd ed.). Technics Publications.

Data Science PM. (2024, December 9). What is CRISP DM? Data Science PM. Retrieved January 6, 2025, from https://www.datascience-pm.com/crisp-dm-2/

Defense Advanced Research Projects Agency. (2024). DPRIVE: Data protection in virtual environments. DARPA. Retrieved December 18, 2024, from https://www.darpa.mil/research/programs/data-protection-in-virtual-environments

Department of Defense. (2016). DOD data management and sharing policy. UAB.edu. Retrieved December 17, 2024, from https://guides.library.uab.edu/rdm/dod

Department of Defense. (2020, September). Executive Summary: DoD Data Strategy. Retrieved December 14, 2024, from https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF

Department of Defense. (2022, March). Summary of the Joint All-Domain Command & Control (JADC2) Strategy. Retrieved January 23, 2025, from https://media.defense.gov/2022/Mar/17/2002958406/-1/-1/1/SUMMARY-OF-THE-JOINT-ALL-DOMAIN-COMMAND-AND-CONTROL-STRATEGY.pdf

Department of Defense Chief Digital and AI Office (CDAO). (2023, June 27). Data, analytics, and artificial intelligence adoption strategy. Retrieved from https://media.defense.gov/2023/Nov/02/2003333300/-1/-1/1/DOD_DATA_ANALYTICS_AI_ADOPTION_STRATEGY.PDF

Department of Defense Chief Digital and Artificial Intelligence Office. (2024, May). Advana 101 briefing deck. Retrieved January 25, 2025, from https://www.acq.osd.mil/asda/dpc/ce/p2p/docs/training-presentations/2024/p2p%202024%20-%20procurement%20analytics%20data%20in%20advana%20part%20i.pdf

Department of Defense Financial Management Transformation. (2022). Department of Defense Financial Management Strategy FY22-26. Retrieved January 25, 2025, from https://comptroller.defense.gov/Portals/45/Documents/DoDFMStrategy/DoD_FM_Strategy.pdf

Desai, K. (2023, October 17). ISO standards for data governance. Medium. Retrieved January 2, 2025, from https://medium.com/data-view-house/why-we-need-iso-standards-for-data-management-f2418898711d

Digital Guardian. (2022, April 27). What is Digital Rights Management (DRM)? (The Definitive Guide). Retrieved December 4, 2024, from https://www.digitalguardian.com/blog/what-digital-rights-management

Duke ScholarWorks. (n.d.). Digital Rights Management (DRM). Retrieved December 9, 2024, from https://scholarworks.duke.edu/copyright-advice/copyright-faq/digital-rights-management-drm/

EDM Council. (n.d.). DCAM framework. Retrieved January 7, 2025, from https://edmcouncil.org/frameworks/dcam/

Enterprise Visibility and Management of Operating and Support Cost (EVAMOSC) Database. (2025, January). How EVAMOSC addresses O&S data challenges. Retrieved from https://evamosc.osd.mil/about.html

Federal Trade Commission. (2009, March 25). Digital Rights Management. Retrieved December 13, 2024, from https://www.ftc.gov/news-events/events/2009/03/digital-rights-management

Go FAIR. (2016). FAIR principles. Retrieved January 6, 2025, from https://www.go-fair.org/fair-principles/

Hero Vired. (2024, September 17). Application of data structures. Retrieved January 16, 2025, from https://herovired.com/learning-hub/blogs/real-time-application-of-data-structures/

Holloway, D. (2022, March 4). Five steps the Defense Department should consider for its data management strategy. C4ISRNet. Retrieved January 2, 2025, from https://www.c4isrnet.com/opinion/2022/03/04/five-steps-the-defense-department-should-consider-for-its-data-management-strategy/

Hußmann, P. (n.d.). Digital Rights Management. Ludwig-Maximilians-Universität München. Retrieved December 5, 2024, from https://www.medien.ifi.lmu.de/lehre/ws0607/mmn/mmn2a.pdf

ISO Policy. (n.d.). United States of America. Retrieved January 8, 2025, from https://policy.iso.org/usa.html

Kiely, T. (2022, August 18). The fundamentals of data structuring. Meltwater. Retrieved January 14, 2025, from https://www.meltwater.com/en/blog/data-structuring

Kosinski, M. (2024, September 6). What is AI data management? IBM. Retrieved January 15, 2025, from https://www.ibm.com/think/topics/ai-data-management

Lin, G. (2021, April 7). Meet Advana: How the Department of Defense solved its data interoperability challenges. Government Technology Insider. Retrieved January 25, 2025, from https://governmenttechnologyinsider.com/meet-advana-how-the-department-of-defense-solved-its-data-interoperability-challenges/

LinkedIn. (n.d.). How do you select and use data management standards and methodologies? LinkedIn. Retrieved January 6, 2025, from https://www.linkedin.com/advice/0/how-do-you-select-use-data-management-standards

Lyon, G. (2001, June). The Internet Marketplace and Digital Rights Management. NIST Advanced Technology Program. Retrieved December 4, 2024, from https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=150308

Manage Engine. (n.d.). What is data tagging? Manage Engine Data Security Plus. Retrieved January 14, 2025, from https://www.manageengine.com/data-security/what-is/data-tagging.html

Marin. (2023, June 19). Data taxonomy best practices. Retrieved January 21, 2025, from https://www.marinsoftware.com/blog/data-taxonomy-best-practices

Merced, A. (2024, March 11). The role of ontologies in data management. LinkedIn. Retrieved January 14, 2025, from https://www.linkedin.com/pulse/role-ontologies-data-management-alex-merced-nfile/

Monsanto, C. (n.d.). What is data management? A complete guide with examples. HubSpot. Retrieved December 16, 2024, from https://blog.hubspot.com/marketing/data-management#types-data-management

National Association of Government Archives and Records Administrators. (2018, November). Digital Rights Management (DRM). Retrieved December 6, 2024, from https://igguru.net/wp-content/uploads/2018/11/NAGARA-Digital-Rights-Management-White-Paper.pdf

National Institute of Standards and Technology. (2024). The NIST Cybersecurity Framework (CSF) 2.0. U.S. Department of Commerce. Retrieved January 4, 2025, from https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf

Office of the Assistant Secretary of Defense. (n.d.). Data analytics. Acquisition. Retrieved January 26, 2025, from https://www.acq.osd.mil/asda/dpc/api/data-analytics.html

Office of the Assistant Secretary of Defense. (n.d.). Data standards - Procurement data standard and other enterprise initiatives. Acquisition. Retrieved January 15, 2025, from https://www.acq.osd.mil/asda/dpc/ce/ds/procurement-data-standard.html

Roncevic, M. (2020, January). DRM and the Law. Library Technology Reports, 10–12. Retrieved December 7, 2024, from https://journals.ala.org/index.php/ltr/article/viewFile/7251/9932

Russakovsky, O., Sadovsky, A., & Ortiz, J. (2004, September). Digital Rights Management. Modern Cryptography: Theory and Applications. Retrieved December 5, 2024, from https://cs.stanford.edu/people/eroberts/courses/soco/projects/2004-05/cryptography/index.html

Shane, R. (2017, February 28). Ontologies: Practical applications. Data Science Central. Retrieved January 21, 2025, from https://www.datasciencecentral.com/ontologies-practical-applications/

Sokolowski, N. (2023, February 15). How is the DoD using data to further the National Defense Strategy? Electrosoft. Retrieved January 26, 2025, from https://electrosoft-inc.com/electroblog/how-dod-using-data-further-national-defense-strategy

Stedman, C. (2023). What are data silos and what problems do they cause? SearchDataManagement. Retrieved January 1, 2025, from https://www.techtarget.com/searchdatamanagement/definition/data-silo

Stedman, C. (2024, May). What is data management and why is it important? Full guide. TechTarget. Retrieved January 28, 2025, from https://www.techtarget.com/searchdatamanagement/definition/data-management

Subramanya, S. R., & Yi, B. K. (2006, April). Digital rights management. IEEE Potentials, 31–34. Retrieved December 2, 2024, from https://www.hit.bme.hu/~jakab/edu/litr/DRM/DRM_Basics_01649008.pdf

Tableau. (n.d.). Guide to data cleaning: Definition, benefits, components, and how to clean your data. Retrieved January 14, 2025, from https://www.tableau.com/learn/articles/what-is-data-cleaning

Tableau. (2024). Data Management: What it is, importance, and challenges. Retrieved December 16, 2024, from https://www.tableau.com/learn/articles/what-is-data-management

Talha, H. (2024, March 25). The best practices for managing and safeguarding government data assets in the digital age. LinkedIn.com. Retrieved December 12, 2024, from https://www.linkedin.com/pulse/best-practices-managing-safeguarding-government-data-assets-haroon-n6eff/

The Data Governor. (2023, April 14). What is data taxonomy with real-life examples? [YouTube Video]. Retrieved January 21, 2025, from https://www.youtube.com/watch?v=06nOHObP6YE

University of California, San Diego Library. (n.d.). Data management best practices. UCSD Library. Retrieved December 17, 2024, from https://library.ucsd.edu/research-and-collections/research-data/plan-and-manage/data-management-best-practices.html

University of Phoenix. (n.d.). What is data management? Definition, examples, best practices. Retrieved December 15, 2024, from https://www.phoenix.edu/blog/what-is-data-management.html

Wharton, B. (2024, October 4). Natural language processing: Transforming large data into strategic business insights. InMoment. Retrieved January 14, 2025, from https://inmoment.com/blog/nlp-natural-language-processing/

Willen, B. (2016). Matthew Green v. DOJ. Retrieved December 11, 2024, from https://www.eff.org/files/2016/07/21/1201_complaint.pdf