Naval Information Warfare Center PACIFIC

TECHNICAL REPORT 3310
JUNE 2023

Enterprise Shared Data Services: A Topical Design Concept Recommendation

David R. Jones
Jeff H. Waters

NIWC Pacific

Naval Information Warfare Center (NIWC) Pacific
San Diego, CA 92152-5001

This page is intentionally blank.

TECHNICAL REPORT 3310
JUNE 2023

Enterprise Shared Data Services: A Topical Design Concept Recommendation

David R. Jones
Jeff H. Waters

NIWC Pacific

Administrative Notes:
This report was approved through the Release of Scientific and Technical Information (RSTI) process in February 2023.

NIWC Pacific
San Diego, California 92152-5001

Patrick McKenna, CAPT, USN,
Commanding Officer

Mike McMillan
Executive Director

ADMINISTRATIVE INFORMATION

The work described in this report was performed by the Command & Control (C2) Technology and Experimentation Division (specifically Branch Codes 53629 and 53622) of the Command & Control and Enterprise Engineering (C2E2) Department, Naval Information Warfare Center (NIWC) Pacific, San Diego, CA.

Released by
Emily Nguyen, Division Head
Command & Control (C2) Technology and Experimentation

Under authority of
Scott Crellin, Department Head
Command & Control and Enterprise Engineering (C2E2)

ACKNOWLEDGMENTS

We would like to acknowledge the support of our colleagues and leadership in our branches, division and department.

Editor: MRM

EXECUTIVE SUMMARY

On September 6, 2022, the Department of the Navy (DON) Chief Information Officer (CIO) signed the Capstone Design Concept for Information Superiority with the enterprise goal to securely move any information from anywhere to anywhere. One of the three objectives for reaching that goal was to adopt enterprise Information Technology (IT) services. The CIO suggested that each of the objectives should support a major design concept, providing a short narrative to educate IT decision-makers on the DON perspective on good design.

This document is a draft design concept addressing recommended guidance for how best to enable Enterprise Shared Data Services (ESDS). Based on previously published Navy data policies and industry best practices, this guidance ensures Navy systems can meet the CIO's goal to securely move any information from anywhere to anywhere, scalably, robustly, maintaining needed context via linked data principles, and with systems built for rapid deployment of new and emerging capability.

OBJECTIVE

Provide guidance on how to integrate and securely deliver visible, accessible, understandable, linked, trustworthy, interoperable, and secure data to worldwide Naval forces, from anywhere to anywhere, on mission timelines to provide decision support.

METHODS

The recommendations were based on principles derived from Navy data policies and crafted using standards and industry best practices.

CONCLUSIONS AND RECOMMENDATIONS

The recommendations include:

1) Apply User-Centered Design

2) Leverage and Contribute to an extensible common Enterprise Data Model

3) Preserve Context by Designing Microservices with Linked Data

4) Confirm and Demonstrate the Enterprise Shared Data Services (ESDS) with Proof-of-Concept Rapid Prototyping

5) Apply User-Driven Testing and Experimentation

6) Leverage Existing Data Services and Data Service Infrastructure

7) Fill Enterprise Data Service Gaps

8) Use a Commercial or Tactical Cloud Base

9) Design Services to Adjust Dynamically to the Enterprise and User Workload

10) Design Development Efforts to Leverage ESDS for Rapid Delivery of Updated Capability

11) Follow and Use the Guidance as Motivational and Confidence-Inspiring

12) Bring in Human Factors (HF) and User-Centered Design (UCD) experts to manage and conduct User Engagements

13) Participate in Communities of Interest to Scale Your Collaboration, Governance and Socialization Efforts

This document provides design and technology patterns to educate and provide specific examples of how to fulfill these recommendations, including:

A) Preserve Data in Context (Linked Data, JSON-LD)

B) Allow for Automation with Machine Understandability (JSON-LD, RDF, OWL)

C) Ease Access to Your Data and Services with a Flexible, Standardized and Scalable Programming Interface (REST, OpenAPI)

D) Prove Viability through Prototyping and a Reference Implementation

This document concludes with a recommendation checklist to enable ease of implementation. In short, this document provides a list of recommendations, techniques, and a checklist process for ensuring that every system is designed to move any information from anywhere to anywhere at any time.

ACRONYMS

ACS	agile core services
AEA	Atomic Energy Act
AI	artificial intelligence
ASW	antisubmarine warfare
C2	command and control
C4ISR	command, control, communications, computers, intelligence, surveillance and reconnaissance
CDO	Chief Data Officer
cm	centimeter(s)
CNO	Chief of Naval Operations
CNWDI	Critical Nuclear Weapon Design Information
COA	course of action
CONUS	continental United States
CUI	Controlled Unclassified Information
CX	customer experience
DAWIA	Defense Acquisition Workforce Improvement Act
DARPA	Defense Advanced Research Project Agency
DASN	Deputy Assistant Secretary of the Navy
DIV	data and information viewpoint
DMO	distributed maritime operations
DoD	Department of Defense
DoDI	DoD Instruction
DoN	Department of Navy
DTM	Directive-Type Memorandum
DVD	digital video disc (also digital versatile disc)
EDM	extensible common enterprise data model
ESDS	Enterprise Shared Data Services
EO	Executive Order
EXDIS	Exclusive Distribution
FDO	Foreign Disclosure Officer
FGI	foreign government information
FOUO	For Official Use Only
FSE	file series exemption
HCS	HUMINT Control System

HF	human factors
HUMINT	human intelligence
IE	DON Information Environment
JSON	javascript object notation
JSON-LD	javascript object notation – linked data
MOC	Maritime Operations Center
NAVWAR	Naval Information Warfare Systems Command
NDS	National Defense Strategy
NIWC Pacific	Naval Information Warfare Center Pacific
OR	operational resilience
OWL	Web Ontology Language
PEO C4I	Program Executive Office for Command, Control, Communications, Computers and Intelligence
PEO MLB	Program Executive Office for Manpower, Logistics and Business Solutions
PPL	people, provenance and lineage
RDF	resource description framework
REST	representational state transfer
SATCOM	satellite communications
SECNAV	Secretary of the Navy
SI	systems integrators
SWAP	space, weight and power
TBM	technology business management
TDC	topical design concept
UAV	unmanned aerial vehicle
UCD	user-centered design
URL	uniform resource locator
VAULTIS	Visible, accessible, understandable, linked, trustworthy, interoperable, and secure
W3C	World Wide Web Consortium
ZTA	zero trust architecture

This page is intentionally blank.

CONTENTS

1. INTRODUCTION: ENTERPRISE SHARED DATA SERVICES

The Enterprise Shared Data Services (ESDS) concept advances the Department of the Navy's (DON) vision for information superiority [1] and details the goal, scope, design principles, objectives, attributes, design and technology patterns that must be employed and adopted to enable, modernize and optimize the DON Information Environment (IE) [2] to drive a competitive advantage as depicted in Figure 1. As participants within the larger ecosystem of DON Enterprise IT Services, ESDS is intended to leverage significant and complementary DON Enterprise IT Services such as authentication, zero trust, and prioritized communications as well as the optimized information environment for cloud for infrastructure services and hosting services.

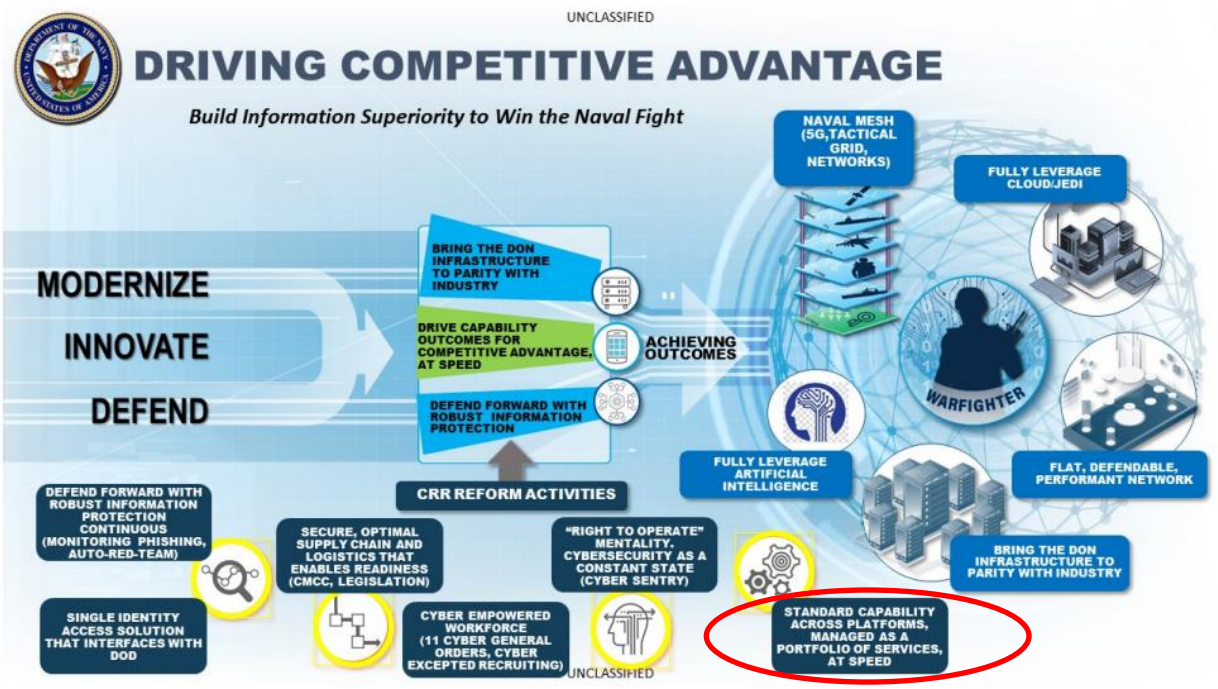

Figure 1. Driving Competitive Advantage overview.

2. GOALS

The goal of ESDS is to integrate and securely deliver visible, accessible, understandable, linked, trustworthy, interoperable, and secure (VAULTIS) [3] data to worldwide Naval forces, from anywhere to anywhere, on mission timelines to provide decision support. The goal of this document and the ESDS Topical Design Concept (TDC), is to provide useful guidance on how to implement the ESDS, based on industry best practices and standards, but independent of specific products or tools. A specific focus is on high-level capabilities to address integration of authoritative data. For clarity and convenience, recommendations in this document are highlighted in italics and prefixed "R:"

3. AUDIENCE

This TDC should be read by anyone with IT responsibilities of any kind (even tangential to the technology itself) regardless of skill level in software development or delivery. The intended audience includes:

- Business analysts with expertise in business processes used to select software applications by functional capabilities
- Program managers whose sole concern is program cost, schedule, and performance
- Contract professionals who write and negotiate contracts for IT products and the Systems Integrators (SI) who bring IT into operation
- Software engineers
- Logisticians
- Cybersecurity professionals
- Executives who must determine appropriate funding

4. SCOPE

The DoD Data Strategy supports the National Defense Strategy (NDS) [4] and digital modernization [5] by providing the overarching vision, focus areas, guiding principles, essential capabilities, and goals necessary to transform the DoN into a data-centric enterprise. The DON Implementation Plan (I-Plan) defines the actions the DON will take to implement the DoD Data Strategy. DON Chief Data Officer published the Naval Data Management Concept of Employment (CONEMP) [6] and the DON Data Architecture [7] to create unity in vision and execution in the DON's organizational and technical implementation of the DoD Data Strategy. ESDS is a critical enabler of this broad data-centric warfighting vision and the guidance in this document includes recommendations which flow directly from and echo the language, recommendations, and decrees in these foundational documents as they apply to ESDS.

The scope of the DON IE, following the Department of Defense Directive (DoDD) 8000.01, Management of the DoD IE, [8] includes: information network infrastructure, enterprise IT service and solutions, National Security Systems, Industrial Control Systems, and embedded computing of wired, wireless, mobile communication and platforms. In terms of the Technology Business Management (TBM) framework, the DON IE capabilities include Data Center, Network, Compute, Storage, Application and Data Platforms, End-User Services, Service Delivery and Operations, Security and Compliance, and IT Management.

The scope of Naval data concerns in this DON IE further incorporates the perspectives, initiatives and recommendations of the DoD Data Strategy, the Creating Data Advantage memorandum, and the emerging Data and Artificial Intelligence (AI) Strategy. [9] The DoD Data Strategy (Sep 30, 2020) envisions a data-centric organization with focus areas on joint all-domain operations, senior leader decision support, and business analytics. The Creating Data Advantage memorandum (May 5, 2021) [10] emphasizes data management and data analytics to create data advantage, treating data as a strategic asset. Meanwhile, the DoD AI Strategy (2018) states the DoD will incorporate AI rapidly, iteratively, and responsibly to enhance military decision-making and operations, and this will include creating a common foundation of shared data, reusable tools, frameworks and standards and cloud and edge services.

5. DESIGN PRINCIPLES

ESDS for integration of authoritative data means modernizing the DON enterprise data handling methods to accomplish master data management, data discovery and retrieval, and data analytics. Design principles serve as guardrails, guideposts and measures, which capture the enterprise intention and guide specific implementations.

The essential overarching enterprise design principles guiding the recommendations in this document and associated guidance incorporated herein are:

- R: VAULTIS - All data will be derived, formatted, integrated and delivered in a manner supporting and enabling VAULTIS. See the DoD Data Strategy defining and supporting VAULTIS [3].
    1) Make Data Visible – Consumers can locate the needed data.
    2) Make Data Accessible – Consumers can retrieve the data.
    3) Make Data Understandable – Consumers can recognize the content, context, and applicability.
    4) Make Data Linked – Consumers can exploit data elements through innate relationships.
    5) Make Data Trustworthy – Consumers can be confident in all aspects of data for decision-making.
    6) Make Data Interoperable – Consumers have a common representation/ comprehension of data.
    7) Make Data Secure – Consumers know that data is protected from unauthorized use/manipulation.

- R: DoD DATA Decrees - All data will be derived, formatted, integrated and delivered in a manner supporting and enabling the five DoD Data Decrees described in the DoD Memo of May 5, 2021, for Creating Data Advantage. [10]
    1) Maximize data sharing and rights for data use: all DoD data is an enterprise resource.
    2) Publish data assets in the DoD federated data catalog along with common interface specifications.
    3) Use automated data interfaces that are externally accessible and machine-readable; ensure interfaces use industry-standard, non-proprietary, preferably open-source, technologies, protocols, and payloads.
    4) Store data in a manner that is platform and environment-agnostic, uncoupled from hardware or software dependencies.
    5) Implement industry best practices for secure authentication, access management, encryption, monitoring, and protection of data at rest, in transit, and in use.

These design principles ensure that data shall no longer be considered a proprietary, ill-defined, poorly tagged, inaccessible, non-machine understandable, and limited resource hidden away in one of many isolated application silos. Instead, data will be shared, accessible, well-defined, and machine-understandable, so that the data can be managed, analyzed by smart systems, and shared to all who appropriately need it to drive warfighting advantage in this information age.

6. OUTCOMES & MEASURES

The outcomes of Customer Experience (CX) and Operational Resilience (OR) will enable the mission's ability to track progress towards this Topical Design Concept's goal to integrate authoritative data. There are and will be many technical and qualitative measures associated with the technology solutions fielded in the DON IE, but, in the end, their success will largely be perceived based on how easy it was for customers to use DON IE solutions to accomplish their missions, and how resilient those solutions were in the face of both deliberate and incidental threats to normal operations. These are not additive outcomes: CX and OR are intrinsic in the design, delivery, sustainment, and operations of all DON IE capabilities.

CX performance measures include:

- Performance Assessment on Selected Tasks and Workflows
  - Human Factors (HF) and User-Centered Design (UCD) experts conduct before and after performance assessment
  - Objective measures include time to complete tasks, accuracy, number of tasks completed, sensor measures of workload
  - Subjective measures include survey assessments by users before and after
- Automated collection of performance and usability data
  - Instrument systems/services to automatically collect user performance data such as time on task, number of completed tasks
  - Instrument users to automatically collect sensor data on stress/workload levels
  - Automate surveys to be conducted periodically
  - Embed feedback, help requests, confidence into systems and services
- Provide the right data to the right place at the right time
  - Number of instances of data authoritativeness adjudications
  - Number of data spillage instances
  - Number of instances where systems are found to be pulling data for non-authoritative sources

OR performance measures are:

- Chaos Monkey-type assessments
- Operator Time spent on systems vice time spent on office products
- Reduction of required help and meeting time
- Number of failures, warning, errors
- Number of Data Sources Ingested
- Number of Data Source Discrepancies Reported / Cleansed
- Essential data handling capabilities for the enterprise Distributed Maritime Operations (DMO) environment

Initial input to methods of measurement will be via the Data Stewards in the form of the operational feedback they receive. (Pending approval)

7. OBJECTIVES

In order to provide enterprise data services for the DON data architecture, one strategy is to organize future activities around three services categories: Data Management and Definition; Data Discovery and Transfer; and Data Transforms and Analytics.

- **Data Management and Definition** services accomplish data identification, registration, adjudication, alignment, update, consistency, and use of authoritative data. The supporting services include services such as cataloging, normalization, mediation, recall, and markup/tagging. Data drives modern warfighting requiring effective knowledge elicitation and common enterprise definitions of warfighter tasks, products, and workflow. These definitions and their machine-understandable data representations are critical enablers of the enterprise data services, supporting modularity, reuse and interoperability. Changing conditions, new missions, and new threats require flexible definitions that can be easily extended but in a well-defined and structured manner. These services are essential for providing authoritative cleansed data for decision making, and they support the design principles of accessibility, trustworthiness, interoperability, linked data and the data decrees of maximizing sharing, automated data interfaces in a standardized manner, authentication, and access management.

- **Data Discovery, Retrieval, Transfer and Storage** services find data and make it available for use across the enterprise. Supporting services include Boolean, faceted, and federated discovery, query management, retrieval, registry, subscription and publishing. These services are essential for providing authoritative, cleansed data for decision making at all echelons, on mission timelines. These services support primarily the design principles of data visibility, accessibility, interoperability, security, and the data decrees to maximize information sharing, publishing of data assets, storing in environment-agnostic manner, access management and protection of data at rest and in transit.

- **Data Transforms and Analytics (including AI and Machine Learning)** services are decision support services that extract meaning from data using specialized algorithms and techniques to organize and interpret data to identify patterns and predict outcomes. These services are currently provided as a subset of enterprise analytic platforms and are emerging as operational and tactical services that complete mindful data reduction providing essential decision support. These services support primarily the design principles of making data visible, understandable, linked, and interoperable; however, they play a special role in the creation of new data and the transformation of that data into something more useful for the warfighter, first information (by applying context) and then knowledge (by applying rules and patterns).

The scope of ESDS encompasses common data handing services, such as mediation, orchestration, publication/subscription, and analytics, provisioned within a common Data Architecture Concept (Figure 2). Data services are componentized, scalable, and configurable to support hosting on all Naval enterprise nodes. Common services and an extensible common Enterprise Data Model (EDM) ensure data-agnostic handling across enterprise nodes, reducing the need for data-specific handling methods.

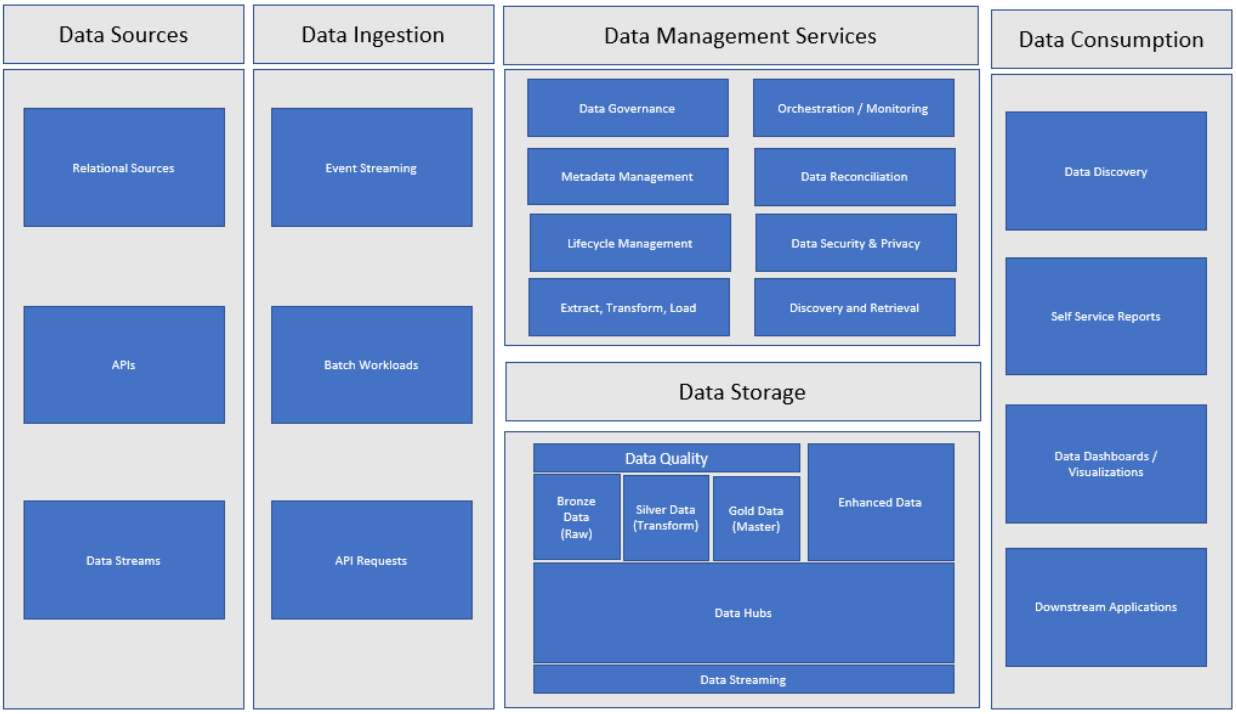

Figure 2. Common Data Architecture Concept.

8. ATTRIBUTES AND RECOMMENDATIONS

Together, the three objectives for this major design concept (Data Management & Definition, Data Discovery and Transfer, and Data Transforms and Analytics) define the strategy for accomplishing integration of authoritative data using enterprise data services; however, if they are not managed correctly, it is possible to pursue these objectives in a conflicting or self-defeating fashion. Before addressing implementation details, consider the following four attributes which characterize the proper mindset for implementing the objectives: Customer-focused; Best-value cost; Dynamic; and Confidence-inspiring. These Attributes can serve as a mental checklist as enterprise data services are designed, delivered, and sustained.

- **Customer-focused:** Customers of the Enterprise Data Services include the DON Information Domains and all data consumers in DON and DoD. Measures of customer performance and satisfaction are described above (Outcomes and Measures). The CX and OR performance measures are focused on the enterprise service capabilities that are essential to our customers. This customer-focus drives the following guidance and required initiatives:
    - R: Apply User-Center-Design (UCD). UCD is a key aspect of our customer interactions and collaboration. UCD applies rigor and science to the DON enterprise through expert-driven user engagements, in-depth knowledge elicitation, and graphical interface model development. UCD defines user tasks, workflows, and products with detail sufficient to enable data modeling and machine understandability. UCD is often the overlooked missing piece in the development of military systems, but this is no longer. UCD is a critical enabler of the Data

Requirements Roadmap because it brings science, clarity, and alignment to the people, systems, and processes of warfighting. [11]

- R: Leverage and Contribute to an extensible common Enterprise Data Model (EDM). An EDM is needed to balance the requirement for common interoperability and reuse with the requirement for extensible tailoring and flexibility to adapt and innovate instantaneously in a dynamic environment. With the foundation of linked data principles, the extensibility can be done in the field by the warfighters themselves. The EDM ensures enterprise stakeholders are all speaking a common language that can grow, evolve, and innovate at the speed of warfighting. The concept of an EDM is discussed and developed in depth in the NAVWAR Navy Enterprise Conceptual Information (DIV-1) document of Dec 7, 2016, and the follow-on NAVWAR Navy Enterprise Data Requirements (DIV-2) Ver 1.0 of March 27, 2017. Initial example representations of the concept and benefits of an EDM in JSON-LD for linked data and OpenAPI for standards-based RESTful API access and extensibility, leveraging CANES Agile Core Services (ACS 3.0), have been explored.
- R: Preserve Context by Designing Microservices (and the foundational EDM and ESDS) with Linked Data. All data exists in context, and it is this context which turns the data into useful information. Rather than ripping the data out of context or attempting to bring in all the external data that represents that context, a best industry practice is to use linked data. [12] The typical pattern is to rely on popular RESTful design approaches, standards such as JSON-LD, and to define an Ontology. This pattern is discussed in more detail below.
- R: Explore, Confirm and Demonstrate the ESDS with Proof-of-Concept Rapid Prototyping (including UCD models of proposed microservices, EDM, and potential extensions and implementations). Prototypes enable advancing beyond paper descriptions to higher fidelity models including source code to ensure the designs on paperwork in real life. Prototypes help demonstrate and socialize the new vision amongst the various stakeholders, including developers and warfighters. Prototyping is a critical enabler for developing and refining an Enterprise Data Requirements Roadmap because it allows us to implement, explore, test and vet our understanding, theories, and ideas prior to investing large costs in time, labor, and money to build the production services. ACS and Overmatch PALS teams are prototyping many of the ESDS concepts using ACS as an exploratory backend. Further information and collaboration is available upon request.
- R: Test the Performance of ESDS-based Microservices in User-driven Testing & Experimentation. User-driven testing and experimentation allows the opportunity to test out the new data representations and architecture in a competitive forum, enabling platform users to try out new concepts to prepare for the fight ahead. Similar to wargaming, experimentation and exercises gives rise to all aspects of warfighting, from situational awareness to planning to course of action (COA) analysis to decision-making to execution and repeating the cycle in the face of a competitive response. In wargaming, people make decisions having consequences for the purposes of education, analysis, experience, and training. Future scenarios can be explored in depth in a manner which can help prepare and avoid conflicts and challenges. Wargaming is a critical enabler for the Data Requirements Roadmap because it is the crucible in which authoritative data must be integrated and new

proof-of-concept data representations and services are measured as an integrated whole in expected warfighting scenarios

- **Best-value cost:** Enterprise Data Services focuses on methods of service development, delivery and support that leverage existing and evolving enterprise organizational capabilities, architecture, and methods of contracting for enterprise services. This approach enables a best-value cost at initiation and ensures continued alignment to best-value cost methods.
    - R: Leverage Existing Data Services and Data Service Infrastructure Where Available. The data services approach includes a range of organizational and technology strategies initiated with efforts to federate existing information domain data hubs and data service capabilities to provide data products across the cross-information domain environment. The approach proceeds to assess, down select, and consolidate to common data services and data handling environments.
    - R: Fill Enterprise Data Service Gaps. We will augment the federation with a focused assessment of available common services, and development to fill gaps in common services. The data ecosystem will move to a common and consistent method of data handling, for all data, across the enterprise, with the goal to eliminate unnecessary data-specific services and costs.
    - R: Use a Commercial or Tactical Cloud Base. We will move to common cloud hosted repositories, leveraging common contracting methods and collaborate with tactical-edge data services to extend data services tactical environments.

  The overall development approach is also a best-value cost approach of using UCD to ensure capabilities are those needed by the warfighter, proceeding with prototyping capabilities to get user feedback, and incorporating needed changes into capabilities early in the development cycle, thus getting the best possible cost investment.

- **Dynamic:** Data services incorporate a dynamic mind-set in the capabilities provided to the end-user, the methods of development, and methods of deployment.
    - R: Design Services to Adjust Dynamically to the Enterprise and User Workload. A key tenet of the enterprise data services is the ability to dynamically adjust to the enterprise environment (e.g., which nodes are active in the decision space) and prioritize data delivery on dynamic mission timelines. To accomplish these tenets, the data services are composable, granular services that can be dynamically delivered/updated to enterprise nodes that may vary widely in Space, Weight, And Power (SWAP) attributes. Data services will be cloud hosted, as possible, and on-prem deployed as needed.
    - R: Design Development Efforts to Leverage ESDS for Rapid Delivery of Updated Capability**.** Our use of UCD to prototype and exercise to obtain quality feedback establishes a dynamic, agile, environment to rapidly respond to changing warfighter needs. The foundational data scheme can be dynamically extended as needed and deployed progressively across the enterprise nodes to meet SWAP constraints while still meeting user data needs.

- **Confidence-inspiring:** The benefits of Enterprise data services and the microservices they enable should be self-evident so as to inspire a new generation of developers to modernize the DoD**.**
    - R: Follow the Guidance in this ESDS-TDC which defines a confidence-inspiring approach and process**.** The guidance in this document is user-focused (UCD-driven), and doesn't stop with modeling, but continues to prototyping and competitive exercises for socialization, vetting, and training. An iterative feedback loop is created to bring the users, and the developers and the designing, along learning, exploring and evolving together.
    - R: Bring in the HF and UCD experts and let them design and lead the User Engagements. This is a gap in most projects and programs and unfortunately, agile scrum approaches tend to exacerbate the problem with a minimalistic and naive approach to use feedback and evolution of requirements. HF and UCD experts should be a permanent part of your team.
    - R: Participate in Communities of Interest to Scale Your Collaboration, Governance and Socialization Efforts. Inspire confidence by implementing organizational governance and management processes that deliver data, to the warfighter, with quality metrics on data; data available ahead of need timelines; compliance and employment of ZTA credentialed monitoring and control of the services and data at the object level; and retaining a PPL chain-of -custody for all access, use and change.

9. DESIGN PATTERNS

ESDS will enable integration of authoritative data and delivery of that data across the enterprise in compliance with the design principles described above by using appropriate design patterns, i.e., general repeatable solutions to commonly occurring problems. Relevant design patterns include:

- R: Enterprise-wide Delivery and Sharing of Data: Deliver the right data to the right place and the right time across the enterprise environment. Virtualized data is disseminated and relocated to meet mission priorities and timeline. Data products are updated/refreshed and staged, across the enterprise, to support tactical use cases. Data movement is dynamically prioritized based on mission priorities and uses prioritized comms to ensure product availability on mission timelines
- R: Source Independence: Critical business processes can be executed across functional domains and systems with secured, timely, accurate and relevant data regardless of the source or location of the data.
- R: Flexible Data Integration for Higher-Level Information and Knowledge: Enable higher-level data by integrating authoritative data sources, and extending common data-agnostic services: useable across the enterprise to discover, access and leverage the entire range of enterprise data using an extensible EDM, enterprise Zero Trust Architecture (ZTA) for data services and data with monitored and controlled visibility and access using person and non-person credentials, metadata registry and management to enable discovery of data and services, multi-cloud, and hardware-agnostic services.
- R: Dynamic Hosting and Scalability: Ensure common data services can be hosted on any enterprise node (enterprise nodes that vary in resources (e.g., Space, Weight, And Power

(SWAP), bandwidth)) and/or can be accessed among nodes for workload sharing. Data Services are composable services that scale to host on any enterprise nodes using an orchestration of fine-grained data services. Deployment of specific capabilities required for enterprise node mission needs. Establish and maintain common scalable data repositories on any enterprise node that any enterprise application can access using common services.

- R: Extensible Common Data Model: Enable reusability, understandability, and interoperability by supporting an extensible common EDM. The EDM can be hosted on each enterprise node and is 1) scalable based on the mission need and resourcing of the enterprise node, 2) includes unique identifiers, and credentialed access (aligned with ZTA) for all data objects, 3) retains People, Provenance, and Lineage (PPL) to support knowledgeable use of data.
- R: Common Lexicons and Metadata: Enterprise data assets, data products, and analytics are published, exposed, and consumed based on common or semantically compatible vocabularies to the greatest degree possible using a common descriptive model.
- R: Reduce Replication: Single purpose copies of datasets are reduced except in the context of their authoritative source and for documented business purposes. Reduce the proliferation of point to-point interfaces. When replicated data does exist, the business purpose, pedigree and lineage, and modifications are fully documented.
- R: Central Metadata Registry: Enterprise metadata is centralized, governed, and integrated at all levels to enable consumers to discover authoritative data and fully understand the business, operational and technical dimensions of enterprise data assets while also being able to preserve their own domain specific language.
- R: Leverage Reference Architectures: A coherent, fully integrated suite of technology enablers and design patterns are employed in a consistent manner to construct all aspects of Enterprise Data Management in a cost-effective manner.
- R: Leverage Shared Services: Enterprise data assets are exposed in a centralized self-service, multitenant analytics, and decision support platform in which descriptions of the available data is exposed through the enterprise metadata management repository.

10. EXAMPLE TECHNOLOGY PATTERNS (EASY BUTTON)

Although this document does not recommend specific products or tools, the guidance provided here does have typical patterns of use in industry, in standards, and in proven techniques. To improve the clarity and usability of the guidance, technology patterns to implement the guidance are provided below. Whereas a design pattern is a general (conceptual) solution to a common problem, a technology pattern is a more specific, technical (industry best practice or open standard) solution to such a problem. In these technical patterns, specific standards and best practices are referenced to enable the solution, not just general concepts. Note that although these patterns mention specific approaches and standards, they still do not specify specific products or tools. These patterns can be implemented with many different tools and in any of a host of programming languages and environments.

Enterprise Data Services and their use to implement VAULTIS data integration in microservices can be explored, prototyped, and demonstrated using a related collection of modern technology patterns critical for effective use of data. The guidance in this document is extended to the recommendation of these patterns:

Preserve Data in Context: Enabling Linked Data in the extensible common Enterprise Data Model (EDM) representation using the W3C standard JSON-LD**.**

Allow for Automation with Machine-Understandability: Clearly defining foundational terms and relationships in a logic-based representation supporting consistency and inferencing using standards such as the Resource Description Framework **(**RDF) with the added context of an associated ontology in the Web Ontology Language (OWL);

Ease Access to Your Data and Services with a Flexible, Standardized and Scalable Programming Interface: Defining the proven and popular RESTful API's in the OpenAPI format (using the EDM represented in JSON-LD); and

Prove Viability through Prototyping and a Reference Implementation: Providing functional test harnesses for testing interfaces (e.g.) as well as prototyping proof-of-concept client front-end and server back-ends to support agile confirmation of new capability effectiveness and efficiency prior to full production development.

Each of these four patterns are described in detail below

- R: Preserve Data in Context (Linked Data, JSON-LD): JavaScript Objection Notation for Linked Data (JSON-LD (https://json-ld.org/)) provides subtle but critical enhancements to the popular JavaScript Object Notation (JSON) to support Linked Data principles (https://www.w3.org/wiki/LinkedData). Linked data principles ensure data is clearly identified, defined, and linked to its context. Linked data maintains context and so supports the design principles of making data linkable, machine-understandable, and accessible. Using JSON-LD, the warfighter terms, tasks, and information products can be represented without sacrificing clarity and context. Using URLs as identifiers and values is one of the tenets of the best practice known as Linked Data. From a practical standpoint, linking data ensures that we never have to strip data away from its context. The problem with JSON from a clear security perspective is that the names of the properties are text strings and not uniquely identifying URLs. This problem is resolved by using the W3C standard JSON-Linked Data (JSON-LD). [13] JSON-LD adds one simple "context" property to JSON, specifying a file (URL) for mapping the property names to URLs. JSON-LD thereby provides the best of both worlds. From a traditional developer perspective, JSON-LD looks just like regular JSON with one additional property. But from a semantic developer perspective, JSON-LD allows the JSON to be converted into RDF and then with the added context of an associated OWL ontology representing the relationships between the terms, the full power of logic-based consistency checking, and inferencing is available. [14] For this reason, this roadmap best represents the tasks, products and workflows in JSON-LD.

- R: Allow for Automation with Machine Understandability (JSON-LD, RDF, OWL): For higher-level automation to support decision-making, data should be represented in a manner that enables machines to clearly identify terms, concepts, objects, and their relationships. JSON-LD is a good first step by mapping JSON property names (text strings) to URLs (unique identifiers). Understandability of data is a key design principle that is necessary for any hope of automation as well as for avoiding confusion or mishandling of data. Identifying terms, values, concepts, objects, and their relationships using URLs is a best practice RESTful and Linked Data technique. The next step to capture the properties and define the relationships of those entities in a standardized machine-understandable format, known as an ontology. When logical characteristics of the relationships are added (e.g., symmetric, reflexive, transitive) then the ontology, if implemented using standards based on description

logic such as the Web Ontology Language (OWL), can support consistency checking and logical inferencing. In short, the ontology represents meaning and knowledge in a manner that allows a machine to do basic reasoning. OWL [15] is one of the semantic standards developed by the World Wide Web Consortium (W3C) along with the Resource Description Framework (RDF) [16]. RDF is an essential modeling technology that allows semantic reasoning between ontological elements to be precisely expressed and captured within multiple contexts. This is critical for interoperability of machine-based and human-based inferencing. Whereas deep learning with neural networks uses impressionistic pattern-matching for its intelligence gained through supervised or unsupervised learning, OWL uses its mathematical foundation with logical inferencing based on a set of well-defined terms and relationships for its intelligence, the other end of the pattern matching spectrum. Together, deep learning and semantic ontologies are a hybrid solution where each technology helps fill the gaps and offsets the weaknesses of the other.

- R: Ease Access to Your Data and Services with a Flexible, Standardized and Scalable Programming Interface (OpenAPI): The Application Programming Interface (API) specifies the method by which a developer can access the tasks, products, workflows, or other services of an application. A popular style of API design is RESTful (https://restfulapi.net/), and a standard for developing RESTful API's is OpenAPI. The use of OpenAPI supports several key design principles including visibility, accessibility, interoperability, and the data decrees to use automated data interfaces externally accessible and machine-readable based on open industry standards, The OpenAPI specification (https://oai.github.io/Documentation/introduction.html) has two main parts: the schema representing the objects such as the tasks, products and workflows, and the routes representing the URLs by which one can access the objects. One nice feature of the OpenAPI schema is that it supports hierarchical reuse of schemas or parts of schemas and so can support building a taxonomic tree of data types that can be represented in JSON-LD, inherited, and extended. A key feature of the EDM is that it should be extensible and the OpenAPI can nicely support this extensibility through its hierarchical schema modeling. Also, the RESTful design of APIs ensures that all important components of the design are accessible via URL. Accessibility of information is one of the important characteristics of modern information warfare.

- R: Prove Viability through Prototyping and a Reference Implementation: After the OpenAPI is defined, there are functional test harnesses that can be utilized to dynamically build a test client front end, such as SwaggerHub. The back end can be auto generated using Swagger tools or by using ongoing communal initiatives such as the integrated knowledge environment supported by Agile Core Services (ACS). The goal is to integrate the microservices to support a reference implementation, i.e., a working proof-of-concept version of the services for exploration, validation, testing, demonstration, and use in exercises and wargaming for end-user feedback. Since the data services will evolve, a continually improving reference implementation will ensure a strong feedback loop to guide the development and ensure the services are meeting the requirements. At this point, the EDM is implemented in its initial form and can be evolved or extended in later iterations. Prototyping is a key pattern for improving, evolving and validating designs, and so prototyping supports all of the design patterns and also facilitates maximizing data sharing as working, provable examples of the design patterns can be shared, not just concepts on paper.

11. MOVING FORWARD

All Topical Design Concepts will be focused, timely, and updated to provide our workforce the up-to-date guidance they need to continue to drive the transformation of the DON Information Environment.

This Enterprise Data Services Topical Design Concept defines the Naval Data Services as a partnership between DON CDO, PEO MLB, and other PEOs. Initial environments will be cloud hosted and will rapidly extend to the tactical edge cloud environment. Initial Enterprise Capabilities will be provided through ADVANA/JUPITER collaboration and service augmentation. Tactical Edge capabilities will be provided in coordination with Over Match Warfighting Domain Services.

12. RECOMMENDATION CHECKLIST

For convenience, the recommendations contained in this document have been assembled into a checklist. Any DoD team or stakeholder developing or using enterprise data, data models or data services can use this checklist to organize discussions to answer how their initiative is enabling and leveraging DoD data policies and Enterprise Shared Data Services and in what ways and to what extent:

Recommended Design Principles and Patterns	**Done? (H/M/L)**
Design Principles	
1. VAULTIS	
2. DoD DATA Decrees	
Attribute: Customer-Focused	
3. Apply User-Center-Design (UCD).	
4. Leverage and Contribute to an extensible common Enterprise Data Model (EDM)	
5. Preserve Context by Designing Microservices (and the foundational EDM and ESDS) with Linked Data	
6. Explore, Confirm and Demonstrate the ESDS with Proof-of-Concept Rapid Prototyping (including UCD models of proposed microservices, EDM, and potential extensions and implementations)	
7. Test the Performance of ESDS-based Microservices in User-driven Testing & Experimentation	

Attribute: Best-cost Value	
8. Leverage Existing Data Services and Data Service Infrastructure Where Available	
9. Fill Enterprise Data Service Gaps	
10. Use a Commercial or Tactical Cloud Base	
Attribute: Dynamic	
11. Design Services to Adjust Dynamically to the Enterprise and User Workload	
12. Design Development Efforts to Leverage ESDS for Rapid Delivery of Updated Capability	
13. Follow the Guidance in this ESDS-TDC which defines a confidence-inspiring approach and process	
14. Bring in the HF and UCD experts and let them design and lead the User Engagements	
15. Participate in Communities of Interest to Scale Your Collaboration, Governance and Socialization Efforts	
Design Patterns	
16. Enterprise-wide Delivery and Sharing of Data	
17. Source Independence	
18. Flexible Data Integration for Higher-Level Information and Knowledge	
19. Dynamic Hosting and Scalability	
20. Extensible Common Data Model	
21. Common Lexicons and Metadata	
22. Reduce Replication	
23. Central Metadata Registry	
24. Leverage Reference Architectures	
25. Leverage Shared Services	
Technology Patterns	
26. Preserve Data in Context (Linked Data, JSON-LD)	

27. Allow for Automation with Machine Understandability (JSON-LD, RDF, OWL)	
28. Ease Access to Your Data and Services with a Flexible, Standardized and Scalable Programming Interface (OpenAPI)	
29. Prove Viability through Prototyping and a Reference Implementation	

13. REFERENCES

[1] DON CIO, "CAPSTONE DESIGN CONCEPT FOR INFORMATION SUPERIORITY", (Sep 2022). https://www.doncio.navy.mil/FileHandler.ashx?id=19037

[2] DON, "DEPARTMENT OF THE NAVY INFORMATION SUPERIORITY VISION", (Feb 2020). https://media.defense.gov/2020/May/18/2002302013/-1/-1/1/DON_INFORMATION_SUPERIORITY.PDF

[3] DoD, "Executive Summary: DoD Data Strategy" (Sep 2020). https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF

[4] DoD, "National Defense Strategy of the United States of America", (Oct 2022). https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.PDF

[5] DoD, "NATIONAL DEFENSE SCIENCE & TECHNOLOGY STRATEGY 2023", (May 2023). https://media.defense.gov/2023/May/09/2003218877/-1/-1/0/NDSTS-FINAL-WEB-VERSION.PDF

[6] UNSECNAV MEMO, "Department of the Navy Actions to Data Advantage", (Jun 2021). https://www.doncio.navy.mil/ContentView.aspx?ID=14828

[7] DON, "Data Architecture", ( ).

[8] DoD, "Management of the Department of Defense Information Enterprise (DoD IE)", (Jul 2017). https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/800001p.pdf

[9] DoD, "Summary of the 2018 Department of Defense Artificial Intelligence Strategy", (Feb 2019). https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

[10] DoD, "Creating Data Advantage", (May 2021). https://media.defense.gov/2021/May/10/2002638551/-1/-1/0/DEPUTY-SECRETARY-OF-DEFENSE-MEMORANDUM.PDF

[11] Cowen, M., Lemon, A., Gill-Hesselgrave, D., "User Centered Design (UCD) Process Description", Technical Report 2061, (Dec 2014). https://apps.dtic.mil/sti/pdfs/ADA615926.pdf

[12] Bizer, Christian, Tom Heath, and Tim Berners-Lee. "Linked data: The story so far." *Semantic services, interoperability and web applications: emerging concepts*. IGI global, 2011. 205-227. https://safe.menlosecurity.com/doc/docview/viewer/docN97100E49CE9B39bf0de1f3d814796e61d2374e43098aa1637f98565a632fbb40c0a77aee7147

[13] W3C Recommendation JSON-LD 1.1, “A JSON-based Serialization for Linked Data” (Jul 2020). https://www.w3.org/TR/json-ld11/

[14] Polleres, Axel, et al. "RDFS and OWL reasoning for linked data." Reasoning Web. Semantic Technologies for Intelligent Data Access: 9th International Summer School 2013, Mannheim, Germany, July 30–August 2, 2013. Proceedings (2013): 91-149. https://safe.menlosecurity.com/doc/docview/viewer/docN6C2FE29287690ff8b9c9182d5de03debcb84cb9bf6c2ba2d211222816d8dcd010bb13520f551

[15] W3C OWL 2, “Web Ontology Language Primer (Second Edition)” (Dec 2012). https://www.w3.org/TR/owl2-primer/

[16] W3C Recommendation RDF 1.1, “Concepts and Abstract Syntax” (Feb 2014). https://www.w3.org/TR/rdf11-concepts/

INITIAL DISTRIBUTION

84310	Technical Library/Archives	(1)
53629	D. Jones	(1)
53622	J. Waters	(1)

Defense Technical Information Center
Fort Belvoir, VA 22060–6218 (1)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-01-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
June 2023	Final	

4. TITLE AND SUBTITLE

Enterprise Shared Data Services:
A Topical Design Concept Recommendation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS

David R. Jones
Jeff H. Waters
NIWC Pacific

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NIWC Pacific
53560 Hull Street
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER

TR-3310

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A: Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This document is a draft design concept addressing recommended guidance for how best to enable Enterprise Shared Data Services (ESDS). Based on previously published Navy data policies and industry best practices, this guidance ensures Navy systems can meet the CIO's goal to securely move any information from anywhere to anywhere, scalably, robustly, maintaining needed context via linked data principles, and with systems built for rapid deployment of new and emerging capability.

15. SUBJECT TERMS

Enterprise Shared Data Services; Capstone Design Concept for Information Superiority; Navy data policies; rapid deployment; secure data; Enterprise Data Model; User-Centered Design; Designing Microservices; Linked Data

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			David R. Jones
U	U	U	SAR	29	19B. TELEPHONE NUMBER *(Include area code)* (619) 767-4781

Naval Information Warfare Center (NIWC) Pacific
San Diego, CA 92152-5001