DoD DATA STRATEGY

2020

FOREWORD

The Department of Defense's (DoD) mission is to provide the military forces needed to deter war and ensure our nation's security. The DoD now recognizes that data is a strategic asset that must be operationalized in order to provide a lethal and effective Joint Force that, combined with our network of allies and partners, sustains American influence and advances shared security and prosperity.

Improving data management will enhance the Department's ability to fight and win wars in an era of great power competition, and it will enable operators and military decision-makers to harness data to capitalize on strategic and tactical opportunities that are currently unavailable. We have a responsibility to gain full value from DoD capabilities and investments, thereby earning the trust of the operational warfighter, the U.S. Congress, and the American people. Embracing new data-driven concepts and leveraging commercial-sector innovations will improve military operations and increase lethality.

To enable this change, the Department is adopting new technologies as part of its Digital Modernization program – from automation to Artificial Intelligence (AI) to 5G-enabled edge devices. However, the success of these efforts depends upon fueling this digital infrastructure in a secure manner with the vast flows of data available from external sources, DoD systems, and connected sensors and platforms. Adversaries are also racing to amass data superiority, and whichever side can better leverage data will gain military advantage. Our ability to fight and win wars requires that we become world leaders in operationalizing and protecting our data resources at speed and scale.

The DoD Data Strategy supports Digital Modernization by providing the overarching vision, guiding principles, essential capabilities, goals, and objectives necessary to navigate this transition and transform into a data-centric enterprise. While opportunities to improve proficiency and efficiency are everywhere, this strategy focuses efforts on Joint Warfighting, Senior Leader Decision Support, and Business Analytics. Success cannot be taken for granted. The responsibility of all DoD leaders is to treat data as a weapon system and manage, secure, and use data for operational effect. The warfighter is counting on us to ensure that the U.S. military remains the most potent and effective fighting force in the world.

David L. Norquist
Deputy Secretary of Defense

TABLE OF CONTENTS

1. INTRODUCTION

The DoD Data Strategy, as a key component of the Department's Digital Modernization program, supports the National Defense Strategy (NDS) by enhancing military effectiveness through access to accurate, timely, and secure data. In addition to combat effectiveness, DoD leaders—including members of the Office of the Secretary of Defense (OSD), the Military Departments, the Office of the Chairman of the Joint Chiefs of Staff (CJCS) and the Joint Staff, Combatant Commands, Defense Agencies, and DoD Field Activities (referred to collectively in this strategy as Components)—require data-driven insights that provide a fair and accurate Department-wide representation of DoD operations and management.

Warfighters at all echelons require tested, secure, seamless access to data across networks, supporting infrastructure, and weapon systems out to the tactical edge. The advanced capabilities provided by DoD's Digital Modernization program depend upon enterprise data management policies, standards, and practices. Sensors and platforms across all domains must be designed, procured, and exercised with open data standards as a key requirement. Survival on the modern battlefield will depend upon leveraging and making connections among data from diverse sources, using analytic tools for superior situational awareness, and coordinating information for disaggregated-precision effects.

This strategy describes the problem and establishes the vision, guiding principles, essential capabilities, goals, and objectives for DoD, relative to data. Figure 1 shows the relationships of these different aspects to one another. The problem statement and scope, stated below, define the Department's first order problem and to whom it applies. The vision statement captures the future state of data. DoD will achieve its vision based on the guiding principles and focused by goals and objectives. Essential capabilities cut across goals and enumerate broad enterprise capabilities.

Figure 1: DoD Data Strategy Framework
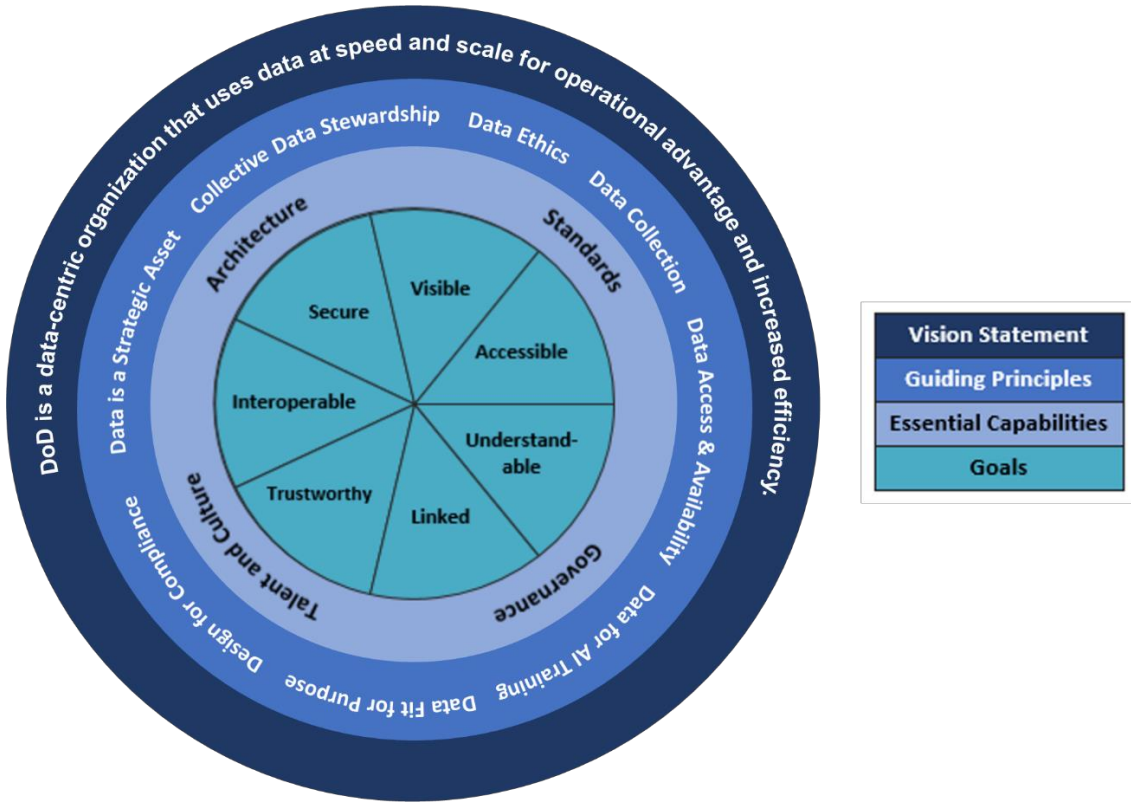
1.1.　Problem Statement

DoD must accelerate its progress towards becoming a data-centric[1] organization. DoD has lacked the enterprise data management[2] to ensure that trusted, critical data is widely available to or accessible by mission commanders, warfighters, decision-makers, and mission partners in a real-time, useable, secure, and linked manner. This limits data-driven decisions and insights, which hinders the execution of swift and appropriate action.

Additionally, DoD software and hardware systems must be designed, procured, tested, upgraded, operated, and sustained with data interoperability as a key requirement. All too often these gaps are bridged with unnecessary human-machine interfaces that introduce complexity, delay, and increased risk of error. This constrains the Department's ability to operate against threats at machine speed across all domains.

DoD also must improve skills in data fields necessary for effective data management. The Department must broaden efforts to assess our current talent, recruit new data experts, and retain our developing force while establishing policies to ensure that data talent is cultivated. We must also spend the time to increase the data acumen resident across the workforce and find optimal ways to promote a culture of data awareness.

1.2. Scope

The DoD Data Strategy applies to the entire Department of Defense and its data, on whichever systems that information resides.

2. VISION AND GUIDING PRINCIPLES

2.1. Vision Statement

*DoD is a data-centric organization that uses data at speed and
scale for operational advantage and increased efficiency.*

---

[1] Data-centric. An environment where data is the primary and permanent asset separated from systems/applications making data available to a broad range of tools and analytics within and across security domains for enrichment and discovery (derived from the IC Data Management Lexicon, January 2020).

[2] Data management. The development and execution of plans, policies, programs, and practices that acquire, control, protect, and enhance the value of data assets throughout the lifecycle (derived from the IC Data Management Lexicon, January 2020).

2.2. Guiding Principles

The Department leverages eight guiding principles to influence the goals, objectives, and essential capabilities in this strategy. These guiding principles are foundational to all data efforts within DoD.

2.2.1. Data is a Strategic Asset

DoD exerts tremendous effort planning and using traditional strategic assets such as personnel, weapon systems, supply chain, and transportation to achieve positive outcomes. In the same manner, data in the DoD is a high-interest commodity and must be leveraged in a way that brings both immediate and lasting military advantage. As DoD shifts to managing its data as a critical part of its overall mission, it gains distinct, strategic advantages over competitors and adversaries alike. These advantages will be reflected in more rapid, better-informed decisions through the use of trustworthy and integrated data.

2.2.2. Collective Data Stewardship

To exploit data fully for decision-making, DoD is defining roles and responsibilities for data stewardship. DoD will assign data stewards, data custodians, and a set of functional data managers to achieve accountability throughout the entire data lifecycle. Data stewards establish policies governing data access, use, protection, quality, and dissemination. Data custodians are responsible for promoting the value of data and enforcing policies, and functional data managers implement the policies and manage day-to-day quality.

2.2.3. Data Ethics

The ethical use of data will be at the forefront of all plans and actions for how data is collected, used, and shared. As the Secretary of Defense stated in his guidance on AI Ethics on February 21, 2020, "*Although technology changes, the Department's commitment to the Constitution, the Law of War, and the highest standards of ethical behavior does not.*" Whether for AI or advanced analytics, ethical principles regarding the responsible use of data remain important, and they will be championed by the DoD CDO and all data and analytics leaders across the Department. Component CDOs will be responsible for promoting a culture of ethical data use supported by oversight mechanisms to identify and promote best practices among the United States and our partners.

2.2.4. Data Collection

Regardless of the data domain, community, or use, the challenge remains the same – to discover and collect data and continuously add value to best inform the decision-maker. Consequently, DoD must enable electronic collection of data at the point of creation and maintain the pedigree of that data at all times. The moment data is created, it should be tagged, stored, and cataloged. When the data is combined or integrated, the resulting product must also be immediately collected, tagged, curated, and appropriately secured. To expedite these processes and to minimize the risk of human error, these steps should be automated to the maximum extent possible.

2.2.5. Enterprise-Wide Data Access and Availability

Closely aligned with data stewardship and collection are the accessibility and availability of data. This is enabled by successful implementation of enterprise capabilities, such as an enterprise cloud; Identity, Credential, and Access Management (ICAM); and associated data-sharing tools. The best technology, processes, and policies will not make this successful if our workforce does not embrace new cultural norms. As such, DoD is making the cultural shift from the need to know (i.e., information withholding) to the responsibility to provide (i.e., information sharing). Making data available across warfighting, intelligence, and business systems is essential to gaining an enterprise-wide view into the daily operations of the Department and absolutely critical to the success of both the National Defense Strategy and the Digital Modernization Strategy. Therefore, it is a fundamental DoD premise that data should be made available for use by all authorized individuals and non-person entities through appropriate mechanisms. To continue this shift, leaders must support this cultural change, set the example, educate their organizations, and enforce data sharing to ensure that the default posture of DoD is to share information broadly. Data sharing should only be restricted when required by law or DoD-wide policy and where security, privacy, or ethical considerations are involved. Furthermore, we must ensure not only that data is protected, but that it is handled properly throughout its lifecycle.

2.2.6. Data for Artificial Intelligence Training

Artificial Intelligence (AI) is a long-term data competency grounded in high-quality training-quality datasets (TQD) that are the pieces of information and associated labels used to build algorithmic models. TQD and the algorithmic models will increasingly become DoD's most valuable digital assets. As DoD modernizes and integrates AI technologies into joint warfighting, generating DoD-wide visibility of and access to these digital assets will be vital in an era of algorithmic warfare. We must also understand that our competitors gain advantage if these assets become compromised. Therefore, the DoD Chief Data Officer (CDO), in partnership with DoD Components, will create a modern governance framework for managing the lifecycle of the algorithm models and associated data that provides protected visibility and responsible brokerage of these digital assets.

2.2.7. Data Fit for Purpose

Data "*fit for purpose*" is quality data that is readily discoverable and understood within the context of its intended use. It should include careful consideration of any ethical concerns in data collection, sharing, use, representation of the information intended, rapid data integration, and minimization of any sources of unintended bias. Customers of DoD data have their own requirements for accessing DoD data, which may or may not align with the purpose or intent of the original data collection. Additionally, in some instances, legislation or a regulation may specify how data is to be used and from which source the data must be consumed. The DoD supports data exploration to enhance analyses that impact decision making.

2.2.8. Design for Compliance

Implementation of IT solutions provides an opportunity to automate the information management lifecycle fully, properly secure data, and maintain end-to-end records management. The Department will make data management and compliance with policies a top priority. Compliance

with required data policies is a critical success factor for continued funding of future warfighting solutions and will be a gate for authorizations to operate.

3. ESSENTIAL CAPABILITIES

Four essential capabilities are needed to accomplish the DoD data goals. These capabilities are not specific to a single goal, but they are necessary to enable achievement of all goals. Only the required capabilities are described in this section. Efforts necessary to achieve the essential capabilities are left to future data strategy implementation plans.

3.1. Architecture

This strategy emphasizes access to data and the capability to adjust requirements in stride with changes in technology and data sources. DoD architecture, enabled by enterprise cloud and other open-architecture capabilities, must allow pivoting on data more rapidly than adversaries are able to adapt. The ability to develop and deploy lightweight applications rapidly and continuously in support of user needs revolutionizes how DoD uses data and leads to a strategic advantage. An agile architectural approach enables incremental value to be delivered by balancing emergent design and intentional architecture. This agile approach allows the architecture of data and systems (even a large solution) to evolve over time, while simultaneously supporting the needs of current users.

3.2. Standards

DoD employs a family of standards that include not only commonly recognized approaches for the management and utilization of data assets, but also proven and successful methods for representing and sharing data. Given the diversity of DoD systems, these standards should be applied at the earliest practical point in the data lifecycle and industry standards for an open data architecture should be used wherever practical. Standards are not an end unto themselves, but rather, they provide value when enabling data and information to be readily and securely utilized and exchanged. Additionally, physical encoding of the data interchange specifications will allow operations in congested and contested environments. Additionally, the DoD CDO will work with the Director, Operational Test and Evaluation (DOT&E), to ensure that data-related material capabilities are tested and evaluated so the effectiveness and suitability of the technologies are known.

3.3. Governance

Data governance provides the principles, policies, processes, frameworks, tools, metrics, and oversight required to effectively manage data at all levels, from creation to disposition. Data governance allows stakeholders to be heard and represented in an organized fashion. For DoD, data governance will be executed at cascading levels, with all issues being resolved at the lowest level possible. Data governance includes localized system decisions affecting data all the way through full records management of critical data assets within the Department. Further, it is essential for data management and records management to be properly implemented throughout the Department.

3.4. Talent and Culture

Moving the Department to a data-centric organization requires a cultural transformation with the DoD workforce at its heart. DoD will continue to evolve its decision-making culture to one soundly based upon data and analytics enabled by technology. A modern, agile, information-advantaged DoD workforce (leaders, service members, civilians, and contractors) will be increasingly empowered to work with data, make data-informed decisions, create evidence-based policies, and implement effectual processes. We must carefully cultivate our data talent to create and sustain these capabilities. We must provide data skill training, establish centers for data engineering excellence, and foster a supportive ecosystem for collaboration among data experts.

4. GOALS AND ENABLING OBJECTIVES

A core tenet of the DoD Data Strategy is the understanding that data is not an IT asset, but an essential and integral part of the mission itself. Data is ubiquitous. DoD weapons platforms, connected devices, sensors, training facilities, test ranges, and business systems generate enormous volumes of data of which all retain and share their data for broader use. It is critical that data be of high quality, accurate, complete, timely, protected, and trustworthy. As such, the Department makes data a strategic asset by establishing the following goals.

DoD data will be:

- **Visible** – Consumers can locate the needed data.
- **Accessible** – Consumers can retrieve the data.
- **Understandable** – Consumers can find descriptions of data to recognize the content, context, and applicability.
- **Linked** – Consumers can exploit complementary data elements through innate relationships.
- **Trustworthy** – Consumers can be confident in all aspects of data for decision-making.
- **Interoperable** – Consumers and producers have a common representation and comprehension of data.
- **Secure** – Consumers know that data is protected from unauthorized use and manipulation.

4.1. Goal: Make Data Visible

The goal of making data visible enables authorized users to discover the existence of data that is of particular interest or value. Data stewards, data custodians, and functional data managers are all responsible and obligated to make their data visible to authorized users by identifying, registering, and exposing data in a way that makes it easily discoverable across the enterprise, and to external partners as appropriate. Moving towards this type of data visibility allows users (person and non-person entities) to discover and rapidly identify who is responsible for specific data assets, the location of data assets, the types of data assets available, and the means of accessing the data assets.

DoD will know it has made progress on making data visible when:

Objective 1: Data is advertised and available for authorized users when and where needed.

Objective 2: DoD implements metadata standards including location and access methods for shared data.

Objective 3: All DoD data sources are catalogued.

Objective 4: DoD implements common services to publish, search, and discover data.

Objective 5: Warfighting and business governance bodies make decisions based on live visualizations of near real-time data.

4.2. Goal: Make Data Accessible

The goal of making data accessible enables authorized users to obtain the data they need when they need it, including having data automatically pushed to interested and authorized users. Data accessibility must comply with Public Law (P.L.) 115-435, the Foundations for Evidence-Based Policymaking Act of 2018. DoD is making data, including warfighting, intelligence, and business data, accessible to authorized users. Accessibility requires that protective mechanisms (e.g., security controls) are in place for credentialed users to ensure that access is permitted in accordance with laws, regulations, and policies.

DoD will know it has made progress on making data accessible when:

Objective 1: Data is accessible through documented standard Application Programming Interfaces (APIs).

Objective 2: Common platforms and services create, retrieve, share, utilize, and manage data.

Objective 3: Data access and sharing is controlled through reusable APIs.

4.3. Goal: Make Data Understandable

Understanding data is critical to enable enhanced, more accurate, and timely decision-making. The inability to aggregate, compare, and truly understand data adversely affects the ability of the Department to react and respond. Without proper context, interpretation and analysis of the data could be flawed, resulting in potentially fatal outcomes. Bringing together business and technology and applying a data-centric approach enable massive amounts of data to be transformed into the insights needed to lead DoD more effectively and efficiently.

DoD will know it has made progress on making data understandable when:

Objective 1: Data is presented in a way that preserves semantic meaning and is expressed in a standardized manner throughout DoD.

Objective 2: DoD utilizes a common data syntax for the same data types and includes semantic metadata with data assets.

Objective 3: Data elements are aligned into a comprehensive data dictionary with a controlled, yet flexible, vocabulary and taxonomy.

Objective 4: Data is baselined and inventoried in comprehensive data catalogs with relevant information on purpose, ownership, points of contact, security, standards, interfaces, limitations, and restrictions on use.

Objective 5: DoD has processes to create, align, implement, and manage business vocabularies, including enterprise standards.

Objective 6: Adaptive, intelligent systems monitor data streams and identify opportunities to transform, combine, or derive new data providing increased insights.

4.4. Goal: Make Data Linked

Data-driven decision-making requires DoD data to be linked such that relationships and dependencies can be uncovered and maintained. Adhering to industry best-practices for open data standards, data catalogs, and metadata tagging, the Department ensures that connections across disparate sources can be made and leveraged for analytics.

DoD will know it has made progress on making data linked when:

Objective 1: DoD implements globally unique identifiers so data can be easily discovered, linked, retrieved, and referenced.

Objective 2: DoD utilizes common metadata standards that allow data to be joined and integrated.

4.5. Goal: Make Data Trustworthy

DoD data requires trust to deliver the needed value to its Service members, civilians, and stakeholders. Lacking confidence in the data may result in less timely decision-making or, consequently, no decision when one is warranted.

DoD will know it has made progress toward making data trustworthy when:

Objective 1: DoD budget requests and the supporting budget process integrate data-focused evidence and Learning Agendas (see P.L. 115-435).

Objective 2: DoD data has protection, lineage, and pedigree metadata bound throughout its lifecycle.

Objective 3: DoD executes data quality management techniques to assess and enhance data quality.

Objective 4: DoD implements master data management for business, intelligence, and warfighting data.

Objective 5: DoD properly tags and maintains all appropriate data and records in accordance with established processes and policies.

4.6. Goal: Make Data Interoperable

Properly exchanging data between systems and maintaining semantic understanding are critical for successful decision-making and joint military operations. Achieving semantic as well as syntactic interoperability using common data formats and machine-to-machine communications accelerates advanced algorithm development and provides a strategic advantage to the Department.

DoD will know it has made progress toward making data interoperable when:

Objective 1: DoD documents and implements data exchange specifications for all systems, including those of coalition partners.

Objective 2: Exchange specifications contain required metadata and convey standardized semantic meaning with the data set.

Objective 3: Public data assets are machine-readable and available for consumption.

Objective 4: DoD rapidly mediates differing data standards and formats without mission-critical loss of fidelity, precision, or accuracy.

Objective 5: DoD develops and promulgates a data-tagging strategy and subsequent implementation plan to enable data interoperability.

4.7. Goal: Make Data Secure

As per the DoD Cyber Risk Reduction Strategy, protecting DoD data while at rest, in motion, and in use (within applications, with analytics, etc.) is a minimum barrier to entry for future combat and weapon systems. Using a disciplined approach to data protection, such as attribute-based access control, across the enterprise allows DoD to maximize the use of data while, at the same time, employing the most stringent security standards to protect the American people.

DoD will know it has made progress toward making data secure when:

Objective 1: Granular privilege management (identity, attributes, permissions, etc.) is implemented to govern the access to, use of, and disposition of data.

Objective 2: Data stewards regularly assess classification criteria and test compliance to prevent security issues resulting from data aggregation.

Objective 3: DoD implements approved standards for security markings, handling restrictions, and records management.

Objective 4: Classification and control markings are defined and implemented; content and record retention rules are developed and implemented.

Objective 5: DoD implements data loss prevention technology to prevent unintended release and disclosure of data.

Objective 6: Only authorized users are able to access and share data.

Objective 7: Access and handling restriction metadata are bound to data in an immutable manner.

Objective 8: Access, use, and disposition of data are fully audited.

5. OPERATIONALIZING THE STRATEGY

Operationalizing the principles in this Strategy is an enduring Department-wide effort. Initial areas for early progress include strengthening data governance, focused engagement in key mission areas, and the execution of detailed Component implementation plans.

5.1. Strengthened Governance

Strengthened data governance will include increased oversight at multiple levels. The Office of the DoD CDO will govern the Department's data management efforts and ensure sustained focus by DoD leaders. The DoD CIO will ensure that data priorities are fully integrated into the DoD Digital Modernization program, ensuring synchronization with DoD's cloud; AI; Command, Control, and Communications (C3); and cybersecurity efforts. The DoD CIO will also promote compliance with CDO guidance via CIO authorities for managing IT investments, issuing DoD policy, and certifying Service/component budgets.

The CDO Council, chaired by the DoD CDO, will serve as the primary venue for collaboration among data officers from across the Department. This body will identify and prioritize data challenges, develop solutions, and oversee policy and data standards of the Department. While working closely with the appropriate governance bodies, members of the CDO Council must also advocate that data considerations be made an integral part of all the Department's requirements, research, procurement, budgeting, and manpower decisions.

5.2. Focus Areas

Data policies and standards alone cannot strengthen data management or improve data quality. They must be continuously informed by feedback from users who consume, produce, manage, and govern data with particular emphasis given to the operational community and warfighter needs. For this reason, the Department must utilize ongoing initiatives in key mission areas to rapidly apply the Strategy's data principles and quickly generate lessons learned. Although data is critical to every DoD mission, initial areas of focus include: Joint All-Domain Operations, business analytics, and senior leader decision support.

Joint All-Domain Operations: As part of the National Defense Strategy's focus on great power competition and conflict, the Secretary has directed the Joint Staff and Military Departments (MILDEPs) to develop new concepts for coordinating military effects in an all-domain fight. The data governance community must closely partner with the Joint All-Domain Command and Control (JADC2) Cross-Functional Team (CFT), the Joint Artificial Intelligence Center (JAIC), and the Deputy CIO for C3 to ensure that we can coordinate information with the tactical edge in a highly contested environment. Clear data standards and interoperability requirements for JADC2 directly support future military readiness. The integrated JADC2 exercises led by the Joint Staff will provide real-world outcomes that will aid in prioritizing data gaps, as will lessons from the Army's work on data design principles and similar efforts by the other MILDEPs.

When new data gaps are identified, the data governance community must work with mission area managers to determine whether changes are needed to hardware; software; tactics, techniques, and procedures; or risk acceptance. The mitigation of many legacy systems is not cost-effective, making it imperative that all future systems are procured with data-interoperability, software upgradability, and cloud-readiness as requirements. The DoD CDO, along with Component CDOs, must also ensure that operational users remain informed of new data-enabled capabilities from the commercial sector and DoD research. Ultimately, DoD's transition to a data-centric organization depends on effective feedback between the data governance and operational communities, and the trusted collaboration this entails.

Senior Leader Decision Support: Senior leaders, including the Deputy Secretary, have directed the development of clear, quantifiable metrics to inform a wide-range of management decisions, such as options for implementing the National Defense Strategy. The data community must support efforts to provide current, decision-quality data along with a platform of tools for analysis and visualization. This approach will accelerate the Department's transition to using live, interactive data in place of static slides to inform strategic outcomes.

Business Analytics: The DoD Comptroller, working with the Chief Management Officer (CMO) and others, is leading an effort to ingest, analyze, and display a wide range of business data; this includes information on budget, procurement, inventory, logistics, and personnel. The data governance community will use the insights from this effort to inform their policies on issues such as authoritative data sources, consistent metadata labeling, standard taxonomies, data provenance, and interfaces (e.g., APIs). This effort could also foster migration to a common data platform for all business analytics across the Department.

5.3. Implementation Plans

To implement this Strategy, Components will develop or convert their extant Data Strategies into measurable Data Strategy Implementation Plans. The CDO Council will support this effort with a common lexicon and common metrics/measures of performance for these plans. This information will be used to identify milestones, track progress, and illustrate data improvements for stakeholders.

6. CONCLUSION

Data underpins digital modernization and is increasingly the fuel of every DoD process, algorithm, and weapon system. The DoD Data Strategy describes an ambitious approach for transforming the Department into a data-driven organization. This requires strong and effective data management coupled with close partnerships with users, particularly warfighters. Every leader must treat data as a weapon system, stewarding data throughout its lifecycle and ensuring it is made available to others. The Department must provide its personnel with the modern data skills and tools to preserve U.S. military advantage in day-to-day competition and ensure that they can prevail in conflict.