

**TITLE:** The Data Steward: The Unsung Hero of AI

**SUB-TITLE:** The roles and responsibilities of data stewards for DoD and Federal Government

**TAGS:** Data Governance, Data Steward, DoD, Data, Federal Government, AI

**Why do we need data stewards?** Data stewards are the hot new role in organizations. They are being talked about across the Internet, at conferences, in the C-Suite, and around the water cooler; but who are they and why do we need them? Before we approach the definition of a data steward, let’s start with why we need them.

DoD and Federal government rapidly adopting Artificial Intelligence (AI) to improve understanding, prediction, and decision making. AI is being applied to every part of Federal and Defense organizations including health, finance, human resources, operations, and security. DoD has even been announcing its intent to move AI to the battlefield. Whether AI is being used to predict disease outbreaks or turn the tide of war it requires accurate, reliable, and trustworthy data. Data stewards are a crucial part in ensuring this happens.

**Data stewards are the foundation of data**

**governance:** The role of the data steward is crucial for implementing data governance in organizations. Data Stewards have a strategic and operational impact across the organization. On a strategic level, data stewards carry out the policies and procedures of data governance and ensure relevance of the data strategy. On an operational level, they engage actively with managing and govern the organization's data assets throughout its lifecycle to ensure data quality, ethical use, and fit for purpose. Some examples of day-to-day activities that data stewards may engage in include furthering organizational data literacy programs, improving data quality, reviewing data classification, informing data policies, encouraging data sharing, improving metadata, maintaining data catalogs, informing data workflows, conducting data monitoring and reporting, and ensuring data security and compliance.

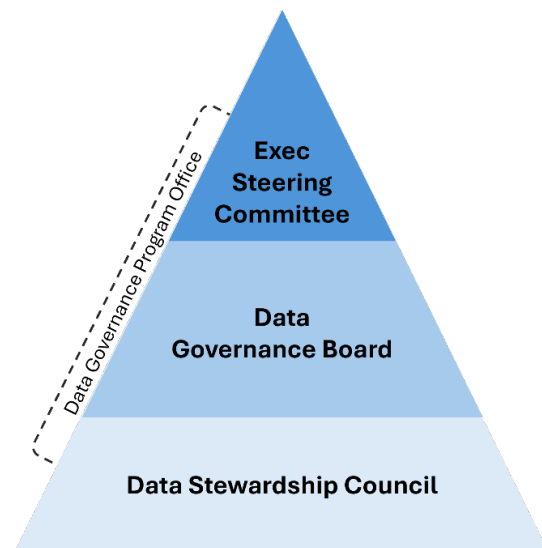


Figure 1: Represents data Stewards as the foundation of data governance

Data stewards also help data governance offices ensure regulatory compliance, improve data control, ensure proper metadata management, encourage data ownership, ease resistance to change, and improve overall organizational data literacy. One resource for supporting data stewards and encouraging active stewardship across large organizations is to implement a data stewardship council. The data stewardship council is the foundation of data governance program and helps provide a platform for organizational data stewards to collaborate and guide the overall stewardship program for the enterprise.

**Types of Data Stewards:** David Plotkin addresses 5 types of data stewards in his book titled “Data Stewards”. Segmenting stewardship roles as David Plotkin suggests may be beneficial for improving data governance programs in large organizations, such as the Federal Government or Department of Defense.

*Lead Data Steward:* The lead data steward runs the Data Stewardship Council and works between the Data Governance Manager and the data stewards to lead day-to-day operations. There is no hard and fast

rule on who the lead data steward should be. Large organizations like DoD or Federal Government may wholly delegate this role to an individual, while smaller organizations may choose to combine this role with the business data steward.

*Business Data Steward:* A Business Data Steward is the key representative in a specific business area that is responsible for quality, use, and meaning of data owned by a business function or domain.

*Project Data Steward:* A Project Data Stewards work with the business data steward. They act as a liaison between the project and the business data stewards to relay data information and questions back and forth.



Figure 2: 5 Types of Data Stewards as described by David Plotkin's in his book "Data Stewards"

*Technical Data Steward:* Where a business data steward is responsible for understanding data domains, technical data stewards focus on the system side of data. Technical data stewards are responsible for understanding the systems, applications, data stores, and technical processes. Technical data stewards help business data stewards understand how the data is created, manipulated, stored, and moved in technical systems.

*Operational Data Steward:* Operational data stewards really accentuate the day-to-day work of data operations. Operational data stewards see early opportunities where improved data quality would benefit certain groups. For example, they may also notice behavior (such as always using a default value) that is damaging to the business.

**Revisiting the quintessential roles of data stewards:** Data Stewards play an extremely important role in many enterprise processes and initiatives that work with data. These roles include inspecting and improving data quality, managing reference data, many of the aspects of master data management (MDM), specifying how data elements should be classified for security and privacy purposes, supporting Quality Assurance (QA), lineage, calculations of process risks, and compliance with privacy regulations.

**Recommendations for implementing data stewardship:** DoD and Federal Government are building and refining data governance programs. This is extremely important to these organizations, ensuring they remain competitive, innovative, and have the best information possible for making strategic decisions for our nation. The success of their data governance programs is crucial to national security and to how our government performs.

At Kinaras Solutions, Inc recognizes that implementing a strong data stewardship program to support data governance and help overcome challenges preventing effective governance is crucial for DoD and Federal Government. Kinaras Solutions, Inc focuses on delivering data governance, data management, cloud, AI, and cyber security solutions to the DoD and Federal Government. We are committed to helping the DoD and Federal Government implement well performing data governance programs and delivering high quality data to their decision-making systems. Through our collaboration and work with the DoD Chief Data and AI Office (CDAO) we have identified several high-value activities for governance programs.

**Establish a data stewardship council:** The data stewardship council is a formal group with the purpose of guiding the overall Data Stewardship effort for the enterprise. A council provides a platform for establishing consistency of the effort and coordination of effort, it facilitates consensus on processes and goals and helps to avoid



confusion. Establishing a centralized council for data stewards is especially useful in large organizations with multi-domain data, such as DoD. The data stewardship council are led the lead data steward (which can also be called the enterprise data steward). Councils are generally attended by business data stewards, which represent their respective data domains for the organization. Business data stewards may also run domain data councils to collaborate with local data stewards on data issues and governance policies. *Kinaras has experience working with data governance offices and can help Federal Government implement data stewardship or data domain councils.*

**Streamline data domains and ensure best practice for data using autonomous data warehouses:**

Data warehouse technologies have long been considered a highly technical and specialized domain. Thanks to technological advancements and cloud technologies, data warehouses have become much more accessible and useable. Two game changing advancements for data warehouses include the autonomous databases and converged database. The autonomous database enables wide adoption of data best practices across enterprises. Because autonomous databases manage themselves any user can quickly and repeatedly provision and use data warehouse anywhere in the organization. Converged database removes the need to deploy and integrate many single purposes databases. It reduces overall data management cost and simplifies data storage by enabling multiple types of data to be stored in one secure location. Some data warehouses have combined these technologies and added modern no-code data interface and streamlined data engagement tools, including all the data lineage, metadata management, user permissions, ETL tools, and data discovery features that a modern organization needs. These super charged data warehouses can be used to streamline complex data operations, create unified data domains, and optimize organizational data fabric. *Kinaras is experienced with autonomous data warehouses and converged databases. We can help Federal Government implement data domain warehouses and leverage best practices for data management.*

**Modernize data use:** Many organizations, particularly those of the Federal Government and DoD managed geographically dispersed data. Data may be stored across systems, databases, data warehouses, or data lakes. While it may be natural for data to be managed locally in geographically dispersed organizations, it often creates risk to data security and quality, increases overall data costs, impedes governance, reduces data visibility, and prevents data access. Data fabric is a powerful solution for enabling data governance in DoD and Federal Government. Implementing data fabric solutions will allow DoD and Federal Government to connect and manage all their data in real time, across different systems and applications. This makes it possible to create a single source of truth, and to use and access data whenever and wherever it is needed. Data fabric does not require data across dispersed systems to be moved, which is beneficial for multi-cloud organizations and organizations or organizations with many individual database systems. This helps improve data access for organizations, while retaining data security and governance. *Kinaras has experience managing data fabric solutions and data architecture for modern enterprises. Kinaras can help the Federal Government implement improve how they manage, access, and use their data through various solutions, including data fabric solutions.*

**Implement modern tools for data governance:** There are many tools available for organizations that are focused on improving data governance. Modern applications include no-code capabilities for data cleaning, data enablement, data quality scoring, data lineage, metadata management, data cataloging, and a host of other things. Data stewards need to be given access to the tools they need to perform the functions of their roles. *Kinaras can recommend available capabilities and assist Federal Government with evaluating and implementing them on their systems.*

**Learn from lighthouses:** Lighthouses serve to warn mariners of dangerous shallows and perilous rocky coasts, and they help guide vessels safely into and out of harbors. In business, the term lighthouses are used to denote an account that has advanced an early adoption of a new technology, service, or solution. They serve as reference models that we can study and learn from. *Kinaras can help Federal Government implement best practices identified from lighthouse organizations.*



Lighthouses for data governance:

- *U.S. Department of Health and Human Services (HHS).* HHS has several data governance resources and initiatives: Data Governance Model, Common Data Use Agreement (DUA) Structure, and HHS Data Governance Board (DGB).
- *United States Geological Survey.* USGS has a robust body of knowledge openly available around the best practices for managing science data, performing data management, understanding data-life cycle, implementing data governance, and building data stewardship.
- *United States Army.* The Army goes rolling along. In this case, the Army has made impressive forward momentum on data governance and data stewardship. The Army has published a plan for building a hierarchy of data stewards and empowering a new role called mission area data officers.
- *U.S. Department of Veteran Affairs.* The VA is the federal government's second largest department after the U.S. Department of Defense and one of the largest data holders. The VA has made significant investments into data governance and AI. The VA's Office of Data Governance and Analytics (DGA) provides a wealth of knowledge and resources around governance and stewardship.
- *Capital One.* There is no industry more heavily regulated than finance. Capital One shares both best practices and general knowledge around data governance in their publication: "A comprehensive guide to data governance".

References:

1. Plotkin, David. *Data Stewardship: An Actionable Guide to Effective Data Management and Data Governance* (pp. 65-66). Elsevier Science.