

Data Stewardship Framework

Social Security Administration

Draft Version 3.0

August 30, 2021

**Table of Contents**

[**Background** 3](#_Toc79140000)

[**Purpose** 3](#_Toc79140001)

[**Data Governance** 3](#_Toc79140002)

[**The Data Stewardship framework** 6](#_Toc79140003)

[**The Data Steward** 7](#_Toc79140004)

[**Critical Success Factors for Data Stewardship** 9](#_Toc79140005)

[**Data Stewardship Roles** 10](#_Toc79140006)

[**Data Stewardship in Action** 13](#_Toc79140007)

[**Selecting a Framework** 16](#_Toc79140008)

[**Glossary** 18](#_Toc79140009)

# **Background**

The Social Security Administration (SSA) began using general-purpose computers in 1955 to handle many of the accounting functions associated with the Social Security program. As the number of federal programs supported by SSA grew, new applications and technologies were developed to collect and store information.

The systems community completed a number of modernization efforts to ensure that the data collected improves the quality of services available to the public. Recent efforts included migrating large amounts of information over to enterprise-wide data structures. Today, the information stored in SSA’s enterprise data warehouse may be accessible to anyone who can write a query, but what safeguards are in place to ensure that the information is accurate, reliable, and secure? How is access granted? Who makes these decisions? These questions can be resolved with the installation of a data governance strategy and a clearly defined data stewardship framework.

# **Purpose**

This document provides an outline for the development of a data governance program at SSA. Data stewardship is the cornerstone of data governance. The framework identifies the roles, responsibilities, and accountabilities of the data stewards. We also identify existing positions at SSA performing data stewardship.

It is important to keep in mind there is no such thing as “*one size fits all*” solution to data governance and stewardship. SSA will have to take a blended approach to be successful. The ability to work within the culture while slowly driving awareness and change is the key to building an effective data stewardship framework.

# **Data Governance**

Data governance is an organizational process and a framework. It requires the systematic creation and enforcement of policies, roles, responsibilities, and procedures. This framework is used by many organizations to establish data ownership and to collaboratively and continuously work to improve data quality. The two key drivers of data governance are:

* Data governance defines the decision-making processes around data.
* Data governance is a strategic, long-term process.

Data governance adds rigor and discipline to the process of managing, using, improving, and protecting organizational information. Effective data governance establishes oversight and common methodologies across the organization. These efforts lead to stronger partnerships, harmonized data, and authoritative analytics that produce relevant, reliable products that help address key business priorities.

**Data Governance Roles**

Without a formal data stewardship framework, organizations must rely on a governance strategy created by the Chief Data Officer and Data Governance Board. They are responsible for the guidance, prioritization and funding of projects and initiatives, as well as the approval of organization-wide data policies and standards. This approach can create inconsistencies as changes in personnel and priorities slow the effort.

Defined in every data governance strategy are the generic roles associated with data management. The three distinct roles (using terms identified by SSA) are the ***data steward, data custodian*** *(owner)*, and ***data manager*** *(custodian).* Briefly summarized, the ***data steward*** is responsible for referencing and aggregating the information, definitions, and any other business needs to simplify the discovery and understanding of these assets. A ***data custodian*** is responsible for the data within their domain in terms of its collection, protection and quality. The ***data manager*** handles server issues, backups, and database maintenance. Their focus is on the technical mastery of database schemas, tables and data integrity.

**The Relationship between Data Governance and Data Stewardship**

Some organizations believe that there is no need to establish a data governance function before or at the same time as the data stewardship initiative.  This is contrary to industry best practices and has shown to be the main reason for the failure of many data stewardship programs.  Data governance is the central component in enterprise data management.  It is the practice that provides planning, oversight, and management of the data and the proper use of data-related resources.

Data governance creates the policies and standards that data stewards implement, and data governance oversees and guides the enforcement of the data stewardship activities.  Without data governance, data stewards would have no organizational policies or global standards to follow. The lack of oversight causes the data stewards to focus on their immediate tactical responsibilities and ignore the need to develop enterprise policies.

Figure 1 illustrates how they work together. The data governance function creates a playbookand focuses on the development of policies and standards for the management information according to data management best practices. The data stewardship function identifies the players and activities needed to implement these policies and ensure that the standards are followed across the organization.  In most large organizations today, the data stewards act as a liaison between the IT department and their business domain, showing the link to the classic definition of the word “steward“: a person who is responsible for managing something on behalf of someone else.

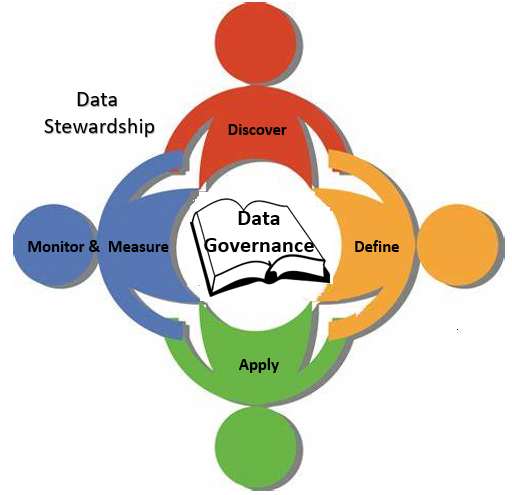


Figure 1- Data Governance and Data Stewardship

After the data governance policies are defined data stewardship provides the basis for the continuance of the program beyond the initial implementation. Data stewardship ensures proper representation across the organization. Data stewardship applies the human factor to data governance that focuses on managing data throughout its life cycle.

It also includes the management and oversight of an organization’s data assets using established data governance practices.  Selected business users ensure that the organization has high-quality data that is accessible to the right user, at the right time, for the right purposes.

Some organizations have created a formal full-time role “Data Steward” and assigned staff to various business domains to serve in that position.  Other organizations have selected their data stewards from staff who remain in their current position (Claims Analyst, Underwriter, Senior Financial Analyst, etc…) and give them the additional responsibilities of business data stewardship.  Regardless of the approach used, data stewards implement the policies and standards created by the data governance strategy.

The Social Security Administration (SSA) is a complex organization with a large number of databases. The agency employs technical data stewards to manage the existing databases. SSA should consider expanding the role of the data steward to support the growing demand for quantitative information across the agency. These new data stewardship roles can be filled with existing employees who possess a combination of four specific skill sets.

1. **Business Knowledge & Experience -** A person who is a subject matter expert within their business domain. They must know what information is collected and why it is needed.
2. **Interpersonal skills -** The data steward must be a good facilitator, communicator, negotiator, and enforcer. They must always look at broader impacts even if they officially only represent one area.
3. **Strong Analytical skills –** such as logical reasoning, critical thinking, research, data analysis, and creativity. The ability to identify priorities, anticipate obstacles, and proactively seek clarity in a group situation.
4. **Technical skills** - Someone who has a general knowledge of the physical implementation of their business domain. They know where to find the data within the computing infrastructure and how it relates to their business domain.

# **The Data Stewardship framework**

This framework provides a general structure and common language for organizing and describing elements of effective data stewardship. The effort is to establish goals, boundaries and principles to allow the agency to implement a data stewardship program that will work for them.

|  |  |
| --- | --- |
| **Strategy and culture** | A strategy that provides a shared vision and clear direction, and a data culture that enables strategic implementation and sustains good data stewardship practices. |
|  |  |
| **Rules and settings** | Legislation, policies, principles, and sanctions providing boundaries and guiding how the data stores should operate. |
|  |  |
| **Roles, responsibilities, accountabilities** | Governance structures, role definitions, expectations, and leadership. |
|  |  |
| **Data capability and quality** | Tools, processes, designs, metadata structures, and platforms for collecting, managing, storing, describing, and sharing data. |
|  |  |
| **People capabilities and literacy** | Skills, knowledge, and services for accessing, managing, analyzing, and communicating data, and insights. |
|  |  |
| **Influence and advocacy** | Effective relationships and networks to endorse, promote, and support good data practices. |
|  |  |
| **Monitoring and assurance** | Assessing environmental trends and developments, measuring stewardship performance and adapting the stewardship framework to respond to changing circumstances or new information |

# **The Data Steward**

A data steward is a person who is responsible for the fitness of data and its context (content and metadata) in a specific domain for an organization. Their job is to identify and curate the quality of the data and transform it into actionable information, and then position it for appropriate use within the organization.They are usually a mid to senior level professional and a subject matter or domain expert. They also understand the organization’s business and have some education in data management concepts.  The candidate should be able to recommend solutions to data challenges and achieve compromises that enhance the use and management of the organization’s data.  Specific competencies include subject area knowledge, data management knowledge (can be acquired through training), analysis skills, and effective communication skills.  Since many data domains cross-organizational boundaries, effective data stewards must possess superior teamwork skills and the ability to see past the confines of their specific component needs to the enterprise’s goals and objectives.

SSA began their data stewardship program in the information technology perspective. Data from earnings, claims processing, and many other systems were stored on files and eventually migrated to databases. To move forward, SSA must make a commitment to expand the data stewardship framework and increase the number of data stewardship roles to include business unit employees, mid-level staff, and/or subject matter experts.

Overall, the expansion of the data stewardship framework will ensure that authoritative data collected provides trusted insights to help resolve critical business problems. Data stewards also improve the quality of the information and enables better decision-making across SSA’s services and operations.

**Data Steward Responsibilities**

Data stewards can have a variety of responsibilities, depending on the organization’s approach to enterprise data management and data governance.  Figure 2 highlights the most common responsibilities performed by the data stewards.

1. ***Business Metadata Management*** – Serve as subject matter experts (SMEs) to answer metadata requests (access and interpretation) from the data governance specialists and other requestors
2. ***Data Management*** – Establish and maintain data quality and integrity in accordance with the policies and procedures laid out by data governance team which include:

* Data exception resolution
* Business rule conformance and data model validation
* Data validation and profiling (in conjunction with data quality team)

1. ***Data Definition Management*** – Develop, enhance, manage, and explain the business data definitions in the data steward’s domain.
2. ***Data Stakeholder and Owner Identification*** – Identify the stakeholder(s) and data owner(s) and their implementation policies.
3. ***Data Usage and Access Management*** – Oversee data access and ensure usage policies are understood and approved, in conjunction with other teams (e.g. data security, enterprise architecture, etc.)
4. ***Data Policies Violation Management*** – Identify and resolve violations of data governance policies and help teams across the enterprise take appropriate corrective action, document and communicate decisions to ensure that policies are followed in the future. Example: Loss/Theft of Personally Identifiable Information (PII)
5. ***Data Change Management*** – Manage changes to data definitions, usage, access, policies, and administration in compliance with data governance standards and best practices.

Figure - Common Data Stewardship Responsibilities

**Common Data Steward Activities**

Performing the data stewardship function includes tasks of a complex nature requiring intimate business knowledge, as well as many tasks that are simple and repetitive that are defined by easily executable business rules. These low-complexity tasks tend to make up the bulk of the volume of the data stewardship activities. In the context of their defined domains, data stewards will:

* Define / describe business terms
* Define data domain values
* Establish and validate data quality rules
* Identify and help resolve data quality issues
* Translate regulatory rules into data polices, rules, and standards
* Help develop data domain business rules & processing requirements
* Define data security requirements
* Promote the use of accepted data definitions and common reference data
* Identify sound data usages and practices
* Determine retention period of data
* Guide and advise others on meaning and proper use of data

# **Critical Factors for Data Stewardship**

Many factors influence the success or failure of a data stewardship initiative. These include:

* ***Establish/Implement a robust enterprise data governance program*** – Data governance is the function that develops the organization’s policies and standards for data and oversees their implementation by data stewards. Organizations that do not have a data governance program will not be able to provide guidance to a data stewardship effort. As a result, any efforts to improve the value of data based on better metadata, higher data quality, organized master data, etc., will be inconsistently implemented, poorly received and ultimately unsuccessful.
* ***Changing the corporate culture*** – A culture that does not embrace the enterprise nature of data stewardship can derail the effort before it starts.  Many stewardship initiatives fail to gain the attention they need or the momentum they deserve because they do not have executive sponsorship or senior management support.  Insufficient funding for enterprise initiatives or a project-based funding model that does not fund any enterprise programs, concerns with lack of commitment of staff resources and a lack of connection to the business goals are some of the reasons there may be resistance to sponsorship for data stewardship framework.  In many organizations, it is essential to have at least one advocate – preferably an execu­tive with vision and organizational authority to ensure that data stewardship is a sustained program across the enterprise.
* ***Providing clearly defined roles and responsibilities for data stewards*** – Organizations that simply assign data stewardship tasks to certain staff members without clearly defining the role and explaining the specific responsibilities of a data steward see poor results from those efforts.  Additionally, those companies that do not invest in training designed for data stewards (foundations of data management, concepts of data governance, task focused training in data definitions and data quality activities) do not realize the sustained benefits that trained data stewards provide.
* ***Lack of measurements and metrics for data steward’s success*** – Measurement is the clearest, simplest way to demonstrate the success of data steward­ship.  The activities of data stewards should be aligned with specific metrics and should be measured regularly. Industry standard measurements include transformation error rates, ratio of data to errors, number of empty values and the amount of unusable (dark) data in a data store.

# **Data Stewardship Roles**

The term data steward generally defines a person who works to ensure the collection and accuracy of an organization’s data. In reality, there can be many different types of data stewards depending on the size and complexity of the organization. Here we have identified the industry standard data stewardship roles as well as the ones currently in use at SSA.

* Executive Sponsor - Chief Data Officer *(****CDO****)*
* Lead Steward
* Business Steward - Product Managers (***PM***) & Subject Matter Expert (***SME***)
* Technical Steward - Enterprise Data Architect (***EDA***)
* Data Manager - Database Administrators & Architects (***DBA, DA***)
* Data Governance Support team *(****CDO support staff****)*

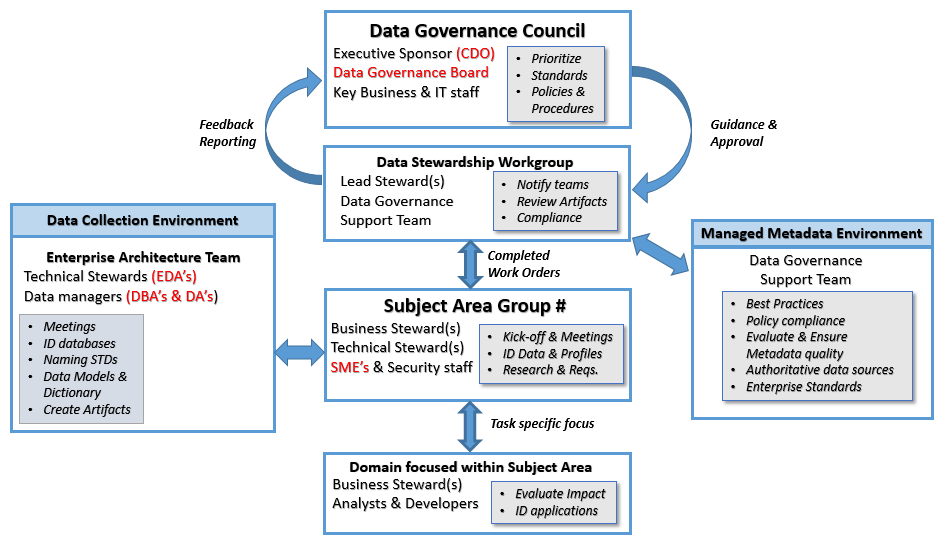
Figure 3 identifies how a traditional data stewardship framework operates and SSA already has pieces (in red) of the framework in place. To move forward we need to evaluate the existing roles and responsibility and determine if this will work for the agency.

Figure 3- Data Steward Roles & Responsibilities

**Executive Sponsor (CDO)**

An Executive Sponsor must facilitate communication, prioritization, funding, conflict resolution, and decision-making across the enterprise. Because data stewardship cuts across a company’s lines of business, it must have executive management support at the highest levels. Executive management involvement is imperative in breaking down the barriers and the organizational silos that exist in all large companies. It is important not to underestimate the political obstacle these silos represent. The political climate is one of the greatest challenges the Executive Sponsor must overcome.

Effective Executive Sponsors do not need to attend every data stewardship meeting, nor do they need to participate in tasks like defining data definitions. Instead, the Executive Sponsor needs to provide the appropriate level of support for their business and technical stewards. Some of the key qualities an Executive Sponsor should have are:

* Someone willing to be an Executive Sponsor
* Someone with an understanding of the value of data as a corporate asset
* Executive Ranking
* High Creditability
* Knowledgeable about the information challenges within the organization
* Willing to challenge the company’s status quo

**Lead Steward**

Lead Stewards are usually needed in large organizations. The Lead Steward is responsible for the overall management of the stewardship function for a particular component or line of business. At SSA, a Lead Steward may represent Earnings or Title 2 and/or Title 16. They manage the stewardship activities of all the data stewards in their business unit. Frequently, they are also responsible for the stewardship of one or more domains within their business unit.

In some organizations, Lead Stewards do not have a direct reporting relationship to line management; rather they serve as leaders of the stewardship team for one of multiple lines of business. Together the Lead Stewards of an organization comprise a data stewardship council or workgroup. They meet to assess the impact of policy changes and co-ordinate how and when they are implemented.

Each Lead Steward must be a highly respected person, who is known throughout the organization. They need to have sound knowledge of both the technical and the business sides of the agency. They must have the ability to manage the politics and conflicts that arise between Business and Technical stewards.

**Business Steward (PM or SME’s)**

Business Stewards are business people who have been charged with the interpretation, enforcement, and conversion of policies into data and metadata requirements – usually in a particular focus area: claims, collections, finance, operations, etc. Depending on the level of the individual steward, he or she may be responsible for advising the organization on governance of categories of data, on definition of the data and its usage, and on the implementation of governance policies through the activities of stewards and data managers. The stewards are responsible for the quality of data and metadata that is part of their functional area, and work to ensure that the governance policies are focused at the enterprise level. Most organizations have several data stewards for each major functional area, while some smaller functional areas may share one steward.

The Business Steward is also responsible for defining the procedures, interpreting policies, data meanings and requirements of the domains assigned to them. They must work with the data on a day-to-day basis and have a strong knowledge of the business requirements, policies and processes. It is essential that they are able to work with other key business components to gain consensus for their organization’s business policies and best practices. At SSA, Business Stewards usually come from the business areas and are usually either Product managers (PM) or Subject Matter Experts (SME).

**Technical Steward (EDA’s)**

The Technical Stewards are members of the organization’s IT department, usually staffed in this role by Enterprise Data Architects (EDA’s). They serve as liaisons between the business areas and IT components, providing expert knowledge of the agency's information assets. These people are the technical resources that manage data naming standards, the creation of data models and identifying where the data fits within SSA’s enterprise architecture.

They assist the Business Stewards by providing help with repository locations, metadata management and data profiling tools. They do not need to possess the policy or domain expertise of the Business Steward or Lead Steward. Their primary focus is to ensure that the data is properly defined, categorized, captured and available to the organization at all times.

**Data managers (DBA’s & DA’s)**

Data managers play a role that is different from a steward. A data manager is the person who implements the data delivery process in concert with the business representatives (the stewards). Data managers provide advice on the technologies used in data management, and enable the user community to access and manipulate (“use”) the data.

In many organizations, data managers reside in the enterprise architecture or database management area. They have developed a deep set of technical data skills that support the data enterprise. They are individuals who are closest to the data and who are responsible for the data and metadata’s quality (accuracy, completeness, validity, etc.). Their goal is to ensure that the data and metadata collected and stored in each application area meet the standards and guidelines developed by the Business Stewards. At SSA, Data Architects (DA’s) and Database Administrators (DBA’s) have been serving in this capacity for many years.

**Data Governance Support Team**

The Data Governance Support Team (DGST) reports to the Chief Data Office and serves as a point of escalation for governance, communication and data quality issues. They are familiar with industry standards and best practices and work closely with business and functional areas to improve the communication, quality and the value of core data assets. The DGST must also provide support to the executive Data Governance Board (DGB) and serve as an independent advisor to assist with regulatory, policy and compliance concerns.

The team includes a small number of mid to senior level staff that have experience building enterprise level data governance programs. They will have strong leadership and communication skills and must have a background in business and technology.

Data Governance Support specialists play a key role in many activities, including:

* Developing data policies, standards, and rules that the enterprise will follow
* Working directly with data stewardship domain groups to ensure that they are adhering to the industry business practices and data governance procedures that have been implemented
* Reviewing and collecting key data governance documents and artifacts like, data dictionaries, data models, data stewardship workflows, meeting notes, etc.

# **Data Stewardship in Action**

Figure 4 provides and overview of what the management structure would look like at SSA when a traditional data governance program is applied. In the design and development portion of the diagram, we specifically highlight the many interactions between the existing data stewards employed by SSA. Teamwork and communication are critical for the successful identification, definition, development and creation of the data artifacts required by the agency.

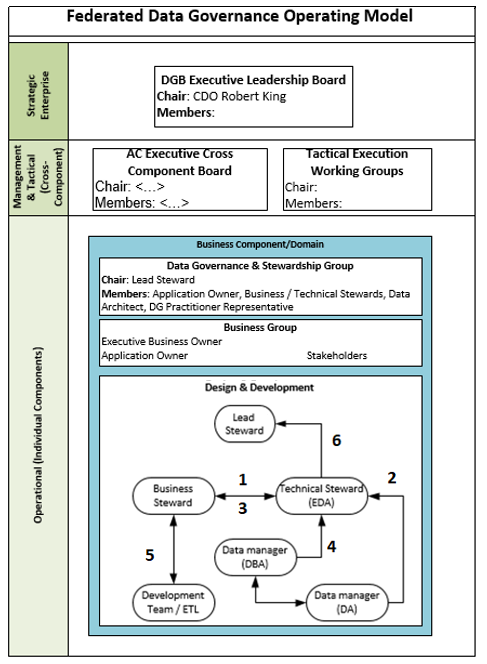


Figure - Data Stewardship Operating Model

1. Business Steward (PM or SME) receives new data requirements and sets up a meeting with the Technical Steward (EDA). They review new requirements and get feedback from the EDA about where the information would fit inside SSA’s current enterprise architecture (new or existing database).
2. The EDA has an internal meeting with the Data managers (DA and/or DBA) to develop data models and naming standards.
3. A second meeting with the Business Steward, to get feedback and make revisions.
4. Another internal meeting (EDA, DBA & DA) to finalize models update dictionaries.
5. Once the data elements are finalized the Business Steward (PM or SME) notifies the development team.
6. EDA notifies Lead Steward and they notify the large systems community to determine impact across systems components.

**Implementation across the framework**

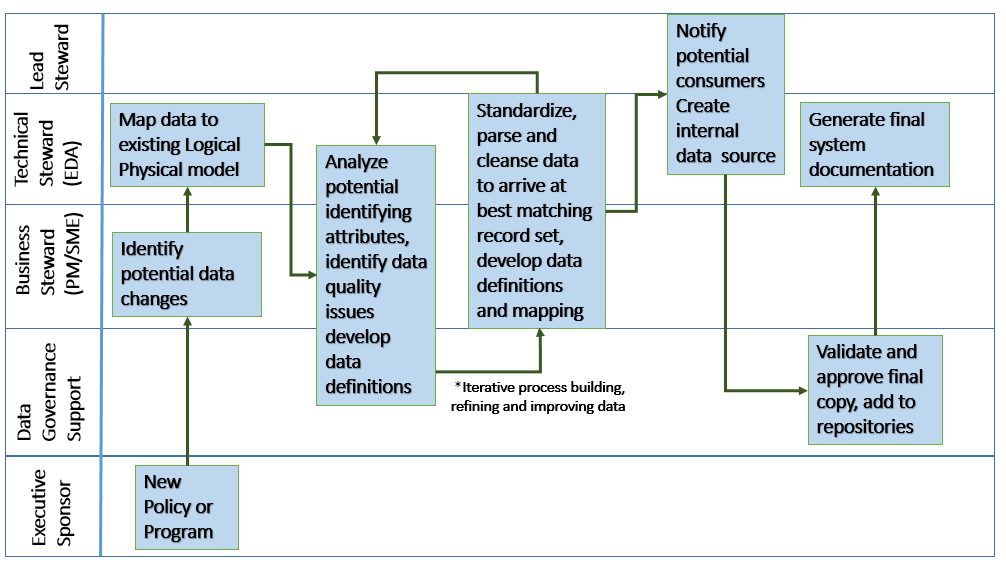
Figure 5 identifies how a data stewardship framework may be implemented to keep each type of data steward involved during the data discovery process. New requirements create the need to collect and store new information. In this example, we identify how the framework is exercised to illustrate stewardship, teamwork and the coordination necessary to implement changes across an organization.

Figure - Workflow across the Framework

When a new policy or program is introduced, it must be sent from the Executive Sponsor to a requirements team for evaluation. The requirements team notifies their Business Stewart for assistance. When new data elements are identified the Business Steward (PM or SME) meets with the Technical Steward (EDA) to discuss the requirements.

Several meetings are conducted that include Business and Technical stewards and may include someone from the Data Governance Support team. At the meetings, many important questions are resolved and work products are created.

* Is the data already being collected in an existing database?
* Is the data added to an existing database or is a new database required?
* What are the data dependencies?
* Naming standards are applied.
* Metadata and data dictionary entries are created/updated
* Data models and data areas are defined
* MI, PII and data security requirements are identified

After all work products are completed, they are passed on to the Data Governance Support Team (DGST) and the Lead Steward is notified. The DGST ensures compliance with governance polices and works with the Technical steward to make sure that the metadata and Enterprise Data warehouse is synchronized with the new information.

# **Selecting a Framework**

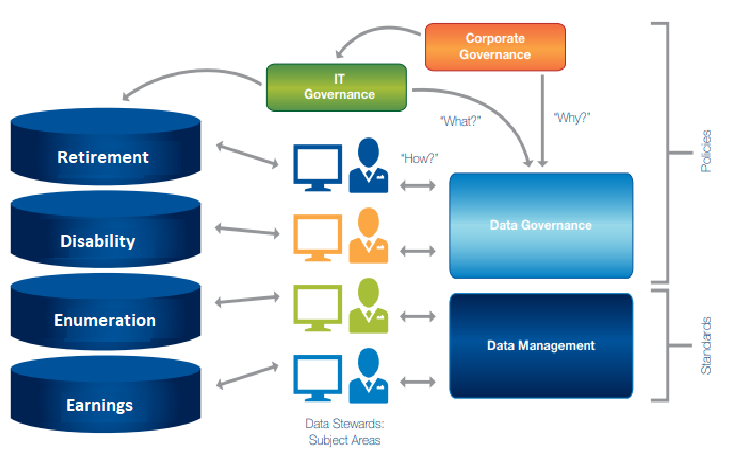
There are many data stewardship frameworks available, but which one will work best for SSA? CGI recommends using the Subject Area/Domain model for data stewardship. Using this model each Business Steward manages a discrete subject area. Therefore, the Business Steward for Retirement processing is different from the Disability Steward, and so on, as illustrated in Figure 6.

Figure 6 - Subject Area/Domain model

Figure 6 shows that a corporate governance process may drive both IT governance and

data governance policies and decision-making. Risk management is part of every organization’s governance framework and includes the rule, “protect personally identifiable information”. The resulting edict will affect IT governance in the form of data, systems and application security policies and will influence how customer data is protected across all domains. Using the Subject Area model, the Business stewards are ultimately accountable for the success and management of their subject area and supporting domains.

In complex environments like SSA, there can be more than one Business steward for each subject area. Depending on the scope and definition of the subject area data stewardship may be further broken down into multiple domains – each with an individual

Business steward. For instance, the “Retirement” area may be broken down into customer, claims, benefit computations and individual awards. One or several stewards may be assigned to each domain depending on the size and complexity of the data contained therein and the breadth of its usage. This approach will work best at SSA because in most cases multiple components share the same data and having defined contact points in each subject area will reduce conflicts, increase accountability and produce higher quality data.

The benefits of a Subject Area/Domain model include:

* Top down approach makes implementation easier
* Ownership boundaries can be clearly defined
* The Business steward’s knowledge of the accompanying business rules and use cases helps expand their knowledge over time (self-fulling)
* Once assigned subject areas can be broken down and domain stewards identified and trained
* Encourages collaboration because subject areas are identified
* The Subject Area model is often easy to pitch: We need someone to manage the Retirement data and most everyone would agree with that.

***Challenges***

**The components that currently perform the data stewardship functions at SSA need to be identified by the subject area and later by domain. From an IT perspective, it should not be difficult. We also have to identify the security teams because they support all areas. The biggest challenge facing SSA in the effort to implement an agency wide data stewardship framework is the need to identify and develop Business Stewards starting with the Subject Matter Experts (SME’s) for each subject area.**

**Once identified these Business Stewards will take on new responsibilities in addition to supporting the existing operational areas. Subject area Business Stewards** will need additional training on data modeling, integration, data configuration and general database concepts. They need to find out which enterprise tools are used in their area and then trained to use them.

# **Glossary**

Refer to: [Data Governance Glossary - Chief Data Officer Space - Confluence](https://confluence.aci.is.cl.ssa.gov/display/CDO/Data+Governance+Glossary)