Data Steward Skill Assessment Survey

Looking to build deeper expertise as a Data Steward?

As part of the Data Stewardship Service, we are offering this *optional* skill assessment to help you get an understanding of the skills that will help succeed in your role as a Data Steward as well as gauge your individual skill level. After completing this assessment you'll receive your results. For the areas in which you'd like to advance your skills, the Data Stewardship Service will provide you with learning path recommendations.

Skill Assessment Overview:
As you move through this assessment, you will first see a set of core data terms. Our team wants to gauge your familiarity with these terms since they are essential to your role. After completing the data terms section, you'll go through **eight skill topic areas** and rate your ability to complete the Data Steward task areas on a scale of **1- 5**, with '1' being the least knowledgeable, to '5' being the most knowledgeable.

Skill Assessment Topic Areas:

1. Data Quality
2. Metadata and Reference Data Management
3. Golden Record Data Management
4. Document and Content Management
5. Data Integration and Interoperability
6. Data Storage and Archival
7. Data Governance
8. Cybersecurity Data Standards

Your answers to this skill assessment survey will be confidential only to you and the Service core team. Our core team will use your results to make customized training recommendations for you so that you can grow your skills as a Data Steward. If you have any questions regarding the survey, please reach out to the CoP Core team at ***datastewardshipservicedevelopment@nasa.onmicrosoft.com***.

Paperwork Reduction Act Statement:

This information collection meets the requirements of 44 U.S.C 3507, as amended by section 2 of the Paperwork Reduction Act of 1995. You do not need to answer these questions unless we display a valid Office of Management and Budget control number. The OMB control number for this information collection is 2700-0153 and it expires on 07/31/2024. We estimate that it will take about 10 minutes to read the instructions, gather the facts, and answer the questions. You

may send comments on our time estimate above to briana.hila@nasa.gov. *Send only comments relating to our time estimate to this address.*

Section 1:

Data Terms

These questions are designed to ensure that as an agency we use the same definition of data terms so that we can work more seamlessly as a data community. If the data term and definition matches with your definition/understanding of the term, select "**Yes**." If it doesn't, select "**No**."

1. **Data Quality:** Data that is accurate, complete, consistent, reliable, up-to-date, and fit for a purpose. Is this the definition you use to define data quality in your day-to-day work.
2. **Data Domain:** A logical grouping of items, or data, of interest to the organization.
3. **Data Catalog:** A software application that creates an inventory of an organization's data assets
4. **Data Interoperability:** The ways in which data is formatted that allow diverse datasets to be merged or aggregated in meaningful ways.
5. **Metadata Management**: Deliberate, structured data about data.
6. **Reference Data**: The data used to classify and reference other data.
7. **Golden Record Management** (also known as **Master Data Management**): A software/tool that logs/manages golden records.
8. **Golden Records** (also known as **Master Data Records**): A single, well-defined version of all the data entities in an organizational ecosystem.
9. **Golden Records** (also known as **Master Data Records**): A single, well-defined version of all the data entities in an organizational ecosystem.
10. **Extract, Transform and Load (ETL)**: A data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system.
11. **Business Glossary**: A collection of data related terms described in clear language that everyone in an organization can understand.
12. **Data Dictionary:** A collection of names, definitions, and attributes about data elements that are being used or captured in a database or information system
13. **Data Integration**: A practice that comprises the practices, architectural techniques and tools for achieving the consistent access and delivery of data across the spectrum of data subject areas and data structure types in the enterprise to meet the data consumption requirements of all applications and business processes.
14. **Data Lineage/Data Flows:** The process of tracking the flow of data over time, providing a clear understanding of where the data originated, how it has changed, and its ultimate destination within the data pipeline.
15. **Data Storage and Archival:** The design, implementation, and support of stored and archived data to maximize its value by managing the availability of data throughout the data lifecycle and ensuring the integrity of data assets

16. **Data Profiling:** A technology for discovering and investigating data quality issues, such as duplication, lack of consistency, and lack of accuracy and completeness.
17. **Centralized System:** Retrieves, collects, and stores data in one place.
18. **Information Assets:** Information relevant to an enterprise's business function, including captured and tacit knowledge of employees, customers or business partners; data and information stored in highly-structured databases; data and information stored in textual form and in less-structured databases such as messages, e-mail, workflow content and spreadsheets; information stored in digital and paper documents; purchased content; and public content from the Internet or other sources.

Selection Options for questions 1-18:
- o Yes
- o No

Section 2:

Data Quality
As a Data Steward, you need to be able to establish data management standards relating to **data quality** in your data domain.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain**: A logical grouping of items, or data, of interest to the organization
- **Data Quality**: Data that is accurate, complete, consistent, reliable, up-to-date, and fit for a purpose

19. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to data quality?

	1: I'm not aware of this task	**2:** I'm aware but need to learn	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task
I can use my data domain expertise to assess data accuracy and validity					
I can write data quality rules to ensure data is complete, consistent, accurate, timely and valid					
I can understand the data quality levels needed so the data can be used for a specific purpose					

I can clean, standardize, and transform data so that it is consistent and suitable for analysis and reporting					

Section 3:

Metadata and Reference Data Management
As a Data Steward, you need to be able to establish data management standards relating to **metadata and reference data management** for your data domain.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization
- **Metadata Management:** Deliberate, structured data about data
- **Reference Data Management:** The data used to classify and reference other data
- **Centralized System:** Retrieves, collects, and stores data in one place
- **Data Catalog:** A software application that creates an inventory of an organization's data assets
- **Business Glossary:** A collection of data related terms described in clear language that everyone in an organization can understand

20. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to metadata and reference data management?

	1: I'm not aware of this task	**2:** I'm aware but need to learn	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task
I can use my data domain expertise in order to assess data accuracy and validity					
I can define metadata based on a human readable, business understanding					
I can define critical data elements for specific use cases within a data asset, and I know who owns the data within a data asset					
I can identify, classify, and define critical data elements					

I can create and follow processes to maintain metadata and reference data criteria and definitions in a centralized system (ex. Data Catalog, Business Glossary, Reference Data Management System)					

Section 4:

Golden Record Data Management
As a Data Steward, you need to be able to establish data management standards relating to **golden record data management** for your data domain.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization
- **Golden Records (**also known as **Master Data Records):** A single, well-defined version of all the data entities in an organizational ecosystem
- **Golden Record Data Management (**also known as **Master Data Management):** A software/tool that logs/manages golden records

21. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to golden record data management?

	1: I'm not aware of this task	**2:** I'm aware but need to learn	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task
I can use my data domain expertise in order to understand the data and determine how to interpret and consolidate data within that domain to create golden records					
I can identify authoritative and relevant sources from which to pull information for golden records					
I can use my data governance					

	1: I'm not aware of this task	2: I'm aware but need to learn	3: I've learned but need coaching	4: I can complete this task on my own	5: I can teach this task
awareness to understand who owns the authoritative data that must be attributed to a golden record					
I can assess the risks related to overmatching and undermatching when consolidating records into golden record					

Section 5:

Document and Content Management
As a Data Steward, you need to be able to establish data management standards relating to **document and content management** for your data domain.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization

22. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to document and content management?

	1: I'm not aware of this task	2: I'm aware but need to learn	3: I've learned but need coaching	4: I can complete this task on my own	5: I can teach this task
I can use my data domain expertise to inventory documents within a data domain					
I can control and organize documents and records through their lifecycle using process creation					

Section 6:

Data Integration and Interoperability
As a Data Steward, you need to be able to coordinate with other data stewards and data custodians to establish data management standards relating to **data integration and interoperability** for your data domain.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization
- **Data Integration:** A practice that comprises the practices, architectural techniques and tools for achieving the consistent access and delivery of data across the spectrum of data subject areas and data structure types in the enterprise to meet the data consumption requirements of all applications and business processes
- **Data Interoperability:** The ways in which data is formatted that allow diverse datasets to be merged or aggregated in meaningful ways

23. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to data integration and interoperability?

	1: I'm not aware of this task	**2:** I'm aware but need to learn	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task
I can use my data domain expertise to assess data accuracy and validity					
I can partner with technical data custodians to provide requirements that improve consistency of data models across systems so data can more easily be integrated					

Section 7:

Data Storage and Archival
As a Data Steward, you need to be able to maintain awareness and support data management standards relating to **data storage and archival.**

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization
- **Data Storage and Archival:** The design, implementation, and support of stored and archived data to maximize its value by managing the availability of data throughout the data lifecycle and ensuring the integrity of data assets

24. As a Data Steward, you need to be able to maintain awareness and support data management standards relating to data storage and archival.

	1: I'm not aware of this task	**2:** I'm aware but need to learn	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task
I can use my data policy awareness to adhere to NASA and federal data storage and archival policies (ex. Records Management policies)					
I can use business acumen to understand the data requirements of my respective NASA office and data domain to determine use cases for data storage and requirements to satisfy those use cases					
I can support storage solutions by partnering with business analysts and technical data custodians to create and improve storage and archival solutions					

Section 8:

Data Governance
As a Data Steward, you need to be able to support compliance efforts by assisting with the implementation of **NASA and Federal data policies**.

To help you complete this section, you may need to reference these data term definitions:

- **Data Domain:** A logical grouping of items, or data, of interest to the organization
- **Data Steward Identification Matrix (DSIM):** The DSIM is a tool to systematically capture the data steward point of contact per data domain and per NASA office

25. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining for each Data Governance competency:

	1: I'm not aware of this task	**2:** I'm aware but need to	**3:** I've learned but need coaching	**4:** I can complete this task on my own	**5:** I can teach this task

		learn			
Communicate Data Standards: I can create content creation for training guides to socialize stewardship related best practices					
Communicate Data Standards: I can lead trainings, conduct coaching sessions to upskill junior data stewards by knowledge sharing stewardship best practices					
Data Policies: I can understand and have detailed familiarity with NASA Data Policies					
Data Policies: I can understand and have detailed familiarity with Federal Data Policies					
Data Steward Identification Matrix: I can maintain the DSIM with my understanding and have awareness of the primary Data Domains for my respective NASA Office					
Data Steward Identification Matrix: I can maintain the DSIM by having awareness of the definition of data stewardship and the typical responsibilities of a Data Steward					

Section 9:

Cybersecurity Data Standards
As a Data Steward, you need to have awareness of **Cybersecurity Data standards**.

26. On a scale of 1-5, how would you rate your ability to complete the following tasks pertaining to cybersecurity data standards?

	1: I'm not aware of	2: I'm aware but	3: I've learned but need	4: I can complete this	5: I can teach this task

	this task	need to learn	coaching	task on my own	
I am policy aware and have detailed familiarity with Federal Cybersecurity Data Policies (ex. NIST recommended security and privacy controls)					
I am policy aware and have detailed familiarity with NASA Cybersecurity Data Policies					
I can implement policy and understand how cybersecurity controls apply to data sharing and accessibility					

Section 10:

Additional Feedback
Your feedback matters - please let us know if there are additional skills for Data Stewards we should include, any concerns or questions you would like to share with the team, or any other general feedback.

    27. Please provide any additional feedback below:
       Answer: _____