GENERAL SERVICES ADMINISTRATION
Washington, DC 20405

CIO P 2142.1, **EXTENDED**
July 1, 2010; May 15, 2020

GSA ORDER

SUBJECT: GSA Information and Data Quality Handbook

1. Purpose. This order issues and transmits Handbook (HB), General Services Administration (GSA) Information and Data Quality Guidelines.

2. Applicability. This order applies to all GSA employees involved in information and data management.

3. Background. The intent of this handbook is to develop a framework for consistent information and dataset management methods. Formally establishing information and data quality principles allows GSA to robustly leverage data for its own business use as well as sharing it with the public. Section 515 of the Treasury and General Government Appropriations Act of 2001 mandates that GSA maximizes the quality, objectivity, utility, and integrity of information it disseminates.

CASEY COLEMAN
Chief Information Officer
Office of the Chief Information Officer

GENERAL TABLE OF CONTENTS

CHAPTER 1 – INFORMATION AND DATA QUALITY GUIDELINES

TABLE OF CONTENTS

1. Purpose. This handbook establishes information and data quality guidelines that can be applied for consistent information sharing and exchange. The scope includes identifying principles that present a high-level description of key components needed for effective enterprise-wide information and data management.

2. Background. Identification of mission-critical information is essential to fulfilling organizational business processes and objectives. Quality information serves as an enabler for cooperation, interoperability, and strategic decision making. Establishing quality guidelines emphasize the development of internal processes that facilitate consistency and help users create value from information and data.

3. Data quality. Data quality entails the effectiveness of the processes and the utility value of the information that is produced. Information quality should include establishing plans for accurately capturing, sharing, distributing, securing, and leveraging data. The enterprise data management framework should provide a logical, model-driven approach to improve quality and overcome barriers to interoperability by:

- Providing a comprehensive model framework that describes common business entities. This framework should also incorporate business rules for data validation, coordination, and integration.
- Employing a common vocabulary that is used to translate among diverse data formats in order to maintain consistency in the meaning of terms.
- Integrating OMB Federal Enterprise Architecture (FEA) Data Reference Model (DRM) 2.0 concepts to enhance organizational capability to discover and reuse federated data.
- Using standards based exchange formats to describe metadata.
- Providing a business modernization blueprint as a methodology for modeling processes and identifying shared services.

The topics mentioned above can be referenced on the enterprise architecture website located at (https://www.gsa.gov/reference/reports/information-quality-guidelinessec-515/data-quality-guidelines).

4. Data stewardship. Each dataset must have a steward responsible for the completeness, accuracy, security, and validity of the data, both at an individual and aggregate level. Stewards should work to guide the appropriate stakeholders in implementing the processes and policies necessary to support quality data.

a. Data stewardship is responsible for ensuring that consistent and cohesive procedures are in place for achieving the following objectives:

- Identifying, administering, and coordinating the alignment of strategic goals to advance the business value of datasets.
- Providing direction to the development and maintenance of datasets and ensure that new and existing datasets are defined in a reliable manner.
- Defining, assigning, and communicating initiatives involving version control, archiving, and dataset decommissioning techniques.
- Continuous quality improvement of datasets to maximize the value of information and increase customer satisfaction.
- Periodic review and update of metadata attributes whereby enabling end-users to search datasets with greater precision.

b. The key supporting stewardship roles include:

- Data Architects - The Data Architect is responsible for the data model and making it business relevant as well as establishing and maintaining naming standards. The Data Architect is responsible for managing and disseminating metadata to facilitate the common understanding of data and encouraging its reuse.
- Data Suppliers - The Data Supplier creates data and maintains its consistency and correctness.

5. Data quality metrics. Managing data quality must include selecting and defining relevant quality measures within the appropriate business context. Table 1.5.1 below provides a list of data quality dimensions that are to be used as a framework for defining rules and quantifying the degree of conformance.

Table 1.5.1 Five Quality Dimensions

Dimension	Description
Accessibility	The degree of ease with which it is possible to locate or obtain data via communications and platform architectures. Availability of sensitive data must be restricted to authorized users.
Completeness	The degree to which requisite data is known and/or recorded. This includes having all facts about the object or event to answer critical business-related questions.
Consistency	The degree to which two data instances provide correlating information about the same underlying object. Are values consistent across data sets? Do interdependent attributes always appropriately reflect their expected consistency?

Dimension	Description
Timeliness	The degree to which data is kept current and available.
Accuracy	The degree to which data correctly reflects the real-world object or event being described. This includes level of precision.

Table 1.5.2 below provides a sample structure for capturing metrics.

Table 1.5.2 Quality Metrics Specification

Factor	Measure	Definition	Threshold
Completeness	Percent Complete	(Available number of values / Total no. of values) x 100	>=99.9%
Timeliness	Percent Timely	Timeliness refers to the delay between the reference period (to which the data pertain) and the date at which the data become available; and the delay between the advertised date and the date at which the data become available (e.g., the actual release date). These aspects are important considerations in assessing quality, as lengthy delays between the reference period and data availability, or between advertised and actual release dates, can have implications for the currency or reliability of the data.	>=99.9%
Accessibility	Average Access Time	The average time required to perform specified data retrieval or manipulation tasks. This time should be determined under realistic conditions (e.g., not only when there is a single connection established to the information system, but when there is concurrent access).	<=3 seconds

6. Information quality. Information quality must be realized through utility, objectivity, integrity, transparency, timeliness, and reproducibility. The information quality principles identified below communicate objectives and the creation of plans for the use and sharing of information across the organization. As identified in Table 1.6, the Information Quality Assessment Process (IQAP), along with its activities and sample work products should be used in analyzing information.

Table 1.6 IQAP Activities and Sample Work Products

Determine Scope	Target Information Assets Define Quality Metrics Cost Analysis	Information Quality Improvement Vision Document
Analyze Information Quality	Assess Logical and/or Physical Model Assess Information Content	Information Quality Analysis Report
Implement Solution	Define Criteria Analyze the Impact	Information Quality Implementation Plan

	Develop the Execution Milestones	

a. Determine Scope - Phase 1 should prioritize target information assets and establish quality metrics.

(1) Target Information Assets - At a minimum, the following activities should be conducted:

- Survey or interview program area stakeholders to elicit information quality issues and requirements;
- Map information quality issues and requirements to IT Strategic Plan goals to prioritize areas for improvement; and
- Develop an information quality improvement plan that includes defining success measures.

(2) Define Quality Metrics - The information quality processes should define metrics applicable to each governed data element. In defining quality metrics the following should be considered:

- The value of the information;
- The cost of errors; and
- The cost of resources to improve the metrics.

(3) Cost Analysis - A comparative analysis of the costs of quality vs. non-quality information should include the following categories:

- Process Failure Costs - Costs incurred when a process cannot be accomplished because of missing, inaccurate, incomplete, invalid, or otherwise poor information;
- Cost of Rework - Costs associated to the time used to reconcile poor or failed information and/or work around processes;
- Opportunity Costs - Indirect costs associated with lost or missed opportunities; and
- Infrastructure Costs - Costs associated with developing and reusing databases and applications.

b. Analyze Information Quality - The goal of Phase 2 should be to create management processes that focus on preventing quality issues.

(1) Assess Logical and/or Physical Model - This activity should involve reviewing the applicable model(s) against business requirements and objectives.

(2) Assess Information Content - Ensure that all parties have a consistent understanding of the information that is being collected and shared.

c. Implement Solution - The following actions should be used when applying the information quality solution in Phase 3:

(1) Define Criteria - The objectives for the solution need to be clearly stated and should correspond with business goals.

(2) Analyze Impact - Evaluate the solution to verify that it accomplishes the desired quality improvement without creating issues.

(3) Develop the Execution Milestones - Identify the schedule, roles and responsibilities, improvement procedures, and lastly the validation methods for applying an information quality solution.

CHAPTER 2 – GSA AND THE DATA QUALITY LIFECYCLE

TABLE OF CONTENTS

1. Purpose. This section outlines the data quality lifecycle that will allow GSA to continually improve its data, and as a result equip GSA to better support its operational, tactical, and strategic activities.

2. Data management roles.

   a. Data management involves the cooperation and communication among the following entities:

   - Data stewards - should work with data owners and managers to continually improve agency data. Responsibilities include defining data governance policies and advising data owners and managers on the implementation of those policies. They should serve as overall coordinators for enterprise data delivery.
   - Data owners - should reside in every business function throughout GSA. Data owners should work with the data stewards to carry the primary responsibility for defining data requirements. Data owners should control access to data as well as oversee changes to data definitions.
   - Data managers - should work closely with the data stewards and data owners to implement data governance policies. There can be multiple data managers throughout the organization. Data managers should be a conduit for improvement ideas.
   - Data users - can be widely characterized as the class of people or processes that utilize data. Data users play a critical role in communicating how data is employed and how it can be improved. Data by itself, without context, has no value, but can become information when a user interprets it.

   b. Sponsorship must include:

   - Chief Information Officer (CIO) - The CIO has responsibility for ensuring that data quality policies and processes align with IT Strategic Plan and business lines. When applicable, the CIO may designate a data management governance workgroup to establish and maintain information quality standards.
   - Heads of Services and Staff Offices (HSSO) - HSSOs are responsible for sponsoring management and technical personnel involved in information and data transparency

initiatives. The information and data derived from HSSOs should serve as an enabler for cooperation, interoperability, and strategic decision making.

3. Data quality lifecycle. In addition to applying the enterprise-wide data quality framework referenced in Chapter 1, GSA should actively engage the principles identified in the data quality lifecycle as an ongoing data management initiative. At the highest level, this process must address the six major components shown in Figure 2.3.

Figure 2.3 Data Quality Improvement Lifecycle

a. Assess - The data quality lifecycle should start with an initial enterprise level or system-by-system assessment of data repositories. The objective should be to analyze the data entries and the data manipulation processes to find the root cause of errors and to highlight improvement opportunities. The output of the assessment should yield an updated data dictionary, clearly defined relationships among data elements, and a roadmap on how the organization will normalize its datasets.

b. Plan - The data owner should determine which improvements have the most far-reaching benefits. After opportunities for improvement have been defined and finalized they should be prioritized, approved, funded, staffed, and scheduled.

c. Execute - The data owner should openly communicate all details of a proposed data quality improvement initiative. This coordination should involve all business users who access the data, the database administrators who are maintaining it, and the developers whose programs have built-in interfaces to the data.

d. Evaluate - The data owner and/or steward should monitor the implemented improvements and determine its effectiveness. They should take into consideration the cost, accuracy, and performance results. If deemed necessary, changes that are problematic should have the flexibility to be reversed with minimal disruption to the organization.

e. Adapt - Data quality improvements that have been tested, verified, and accepted must be announced to the entire organization before turning them into new standards, guidelines, or procedures.

g. Educate - The final phase is to disseminate information about the data quality improvements that have been implemented. Depending on the scope of the change, education can be accomplished through the organization's intranet, an internal newsletter, or a broadcasted e-mail to all stakeholders.

4. Data security. Throughout its lifecycle, data must be protected in a manner that is defined by the policies and procedures of GSA. The Information Security Office (IS) should be contacted directly if at any point there is concern or questions regarding any aspect of GSA data, to include but not limited to topics in access control, processing, or handling instructions.

5. Data sharing. All data sharing efforts within GSA must meet agency-wide statutory, regulatory, and security mandates.

a. Internal - To maximize its use of information, GSA must enable the frictionless-flow of data within the agency. Data should be made available in a timely and responsive manner in support of the following strategic goals:

- uncovering business opportunities;
- improving operational efficiency and effectiveness; and
- informing GSA's key stakeholders.

b. External - GSA's default position will be to share information openly with the public while meeting existing requirements that protect the release of inappropriate data. The example guidelines illustrated in Appendix A identify criteria that will allow GSA to fully participate in DATA.gov submissions. As other external data sharing programs are formalized, additional processes must be customized according to unique considerations.

APPENDIX A

DATA.gov Submissions

1. Background. DATA.gov is an initiative to allow the public to easily find, download, and use datasets that are generated and held by the Federal government. Improving access to GSA data can help foster innovation and fuel the knowledge economy as well as increase transparency. DATA.gov will enable the public to participate in research and discovery by providing data from which they can build applications and conduct independent analyses.

2. Purpose. This section identifies GSA specific procedures and guidelines that anyone involved in releasing datasets to DATA.gov (e.g. program offices, data owners/stewards) should follow.

3. Metadata. The data steward must describe a dataset according to attributes formalized by the DATA.gov Program Management Office (PMO). Metadata will provide information about the context of the data collection, dataset completeness, and other factors that might influence the utility of the data for a specific purpose. The most critical elements of metadata include:

- data descriptions;
- keywords;
- data sources;
- URLs of technical documentation; and
- security considerations.

The data steward should think both broadly and specifically when selecting keywords; the robustness of the text-based search capability will determine the extent to which users can find the data in which they are interested. The most current metadata template may be obtained from the GSA Point of Contact (POC) or be directly entered through the DATA.gov Data Management System (DMS) (https://www.data.gov/).

4. Impact assessment. The primary objective of an impact assessment is to identify the potential consequences of releasing datasets to DATA.gov. The data steward must ensure that no adverse effects will occur as a result of making its data public. Where applicable, the impact assessment may consist of two parts: technical analysis followed by risk management. Program offices should continue to follow their regular routine for making information or data public to include however many levels of approval they deem necessary. If a data steward is not confident regarding the components of an impact assessment they should contact the respective organization responsible for establishing the policy and/or regulation.

The GSA guidelines below should be thoroughly examined to determine the suitability of sharing candidate datasets:

a. Security - (Security Categorization, Risk Assessment, Certification and Accreditation) (CIO IT Security 06-30) - 10/3/2019 (Managing Enterprise Cybersecurity Risk [CIO-IT Security 06-30 Rev-25] [PDF - 1 MB].

b. Privacy – CIO 2231.1 GSA Data Release Policy (https://www.gsa.gov/directive/gsa-data-release-policy).

c. Freedom of Information Act (FOIA) - Exemptions (https://www.gsa.gov/reference/freedom-of-information-act-foia).

d. Legal - Laws and Legal Matters
https://www.gsa.gov/policy-regulations/policy/information-integrity-and-access/it-security-procedural-guides.

5. Data policy statements. Program offices and data stewards should adhere to the default data policy guidelines described below prior to releasing any data to the public.

a. Public Information - Datasets accessed through DATA.gov are confined to public information and must not contain National Security information as defined by statute and/or Executive Order, or other information/data that is protected by other statute, practice, or legal precedent.

b. Security - Information accessed through DATA.gov must be in compliance with the required confidentiality, integrity, and availability controls mandated by Federal Information Processing Standard (FIPS) 199 as promulgated by the National Institute of Standards and Technology (NIST) and the associated NIST publications supporting the Certification and Accreditation (C&A) process.

c. Privacy - Information accessed through DATA.gov must be in compliance with current privacy requirements including OMB Circular A-130. In particular, GSA is responsible for ensuring that the datasets accessed through DATA.gov have any required Privacy Impact Assessments or System of Records Notices (SORN) easily available on GSA.gov. Under no circumstances, should data contain Personal Identifiable Information (PII).

d. Data Quality and Retention - Information accessed through DATA.gov is subject to the Information Quality Act (P.L. 106-554). For data accessed through DATA.gov, GSA must confirm that the data being provided meets GSA's Information and Data Quality Guidelines.

As the authoritative source of the information, GSA retains version control of GSA datasets accessed through DATA.gov in compliance with record retention requirements outlined by the National Archives and Records Administration (NARA).

e. Secondary Use - Data accessed through DATA.gov do not, and should not, include controls over its end use. However, as the data owner or authoritative source for the data, GSA must retain version control of datasets accessed. Once the data have been downloaded, GSA cannot vouch for quality and timeliness. Furthermore, GSA cannot vouch for any analyses conducted with GSA data retrieved from DATA.gov.

6. Office of Information and Regulatory Affairs (OIRA) guidance.

a. GSA retains responsibility as the authoritative source of data, including corrections and updates.

b. GSA retains responsibility for protection of personally identifiable information and records retention.

c. GSA program office attests:

- The dataset is in compliance with applicable privacy, confidentiality, and other relevant statutes;

- The dataset is in compliance with agency Information and Data Quality Guidelines;
- Exposure support; and
- Performance - site capability.

7. Information and data dissemination criteria. In an effort to make data dissemination consistent, as well to improve compliance with existing statutory responsibilities, the OCIO and its designated governance working group have created a dataset quality checklist. Table A-7 should be referenced to ensure that datasets being released outside the agency conform to GSA's criteria.

Table A-7 GSA's DATA.GOV Dataset Checklist

DATA.gov Submission Checklist

Dataset Name & Description:

---

a) Data Source –

---

b) System / Application –

---

c) Target Audience –

High-value* [Y / N]:

☐ *Increase agency accountability and responsiveness*
☐ *Improve public knowledge of the agency and its operations*
☐ *Further the core mission of the agency*
☐ *Create or expands economic opportunity*
☐ *Respond to need and demand as identified through public consultation*
☐ *Other*

If YES, please provide a brief explanation geared to the citizen reader of why it is high-value:

1	☐	The dataset is public information and does not contain National Security or other information/data protected by statute, Agency practice, legal precedent, or otherwise restricted by GSA.
2	☐	The dataset complies with required confidentiality, integrity, and availability controls for GSA, thereby adhering to NIST and OMB guidance.
3	☐	The dataset is in compliance with GSA and OMB privacy requirements.
4	☐	The data owner (*signature below*) certifies the dataset meets GSA's Information Quality and Data Guidelines (CIO P 2142.1) to include the following components defined in Table 1-5-1 of the Handbook. ▪ Accessibility ▪ Completeness ▪ Consistency ▪ Timeliness ▪ Accuracy

		Where applicable please identify means, mechanisms, or persons that conducted/attest to the assessment:
5	☐	The data owner and/or program office is the authoritative* source for the data and manages versioning and record retention requirements.
6	☐	The dataset does not include controls over its end use. In the metadata for the dataset, the Agency citation should note that the data was obtained from DATA.gov and that the Federal Government cannot vouch for its analyses after being retrieved from DATA.gov.
7	☐	The dataset is a product of the Federal Government, or the government has unrestricted rights of use. The dataset must be suitable for listing or downloading through any of the DATA.gov catalogs. The "Raw" Data Catalog provides an instant download of machine readable, platform-independent datasets, while the "Tools" Catalog provides hyperlinks to webpages that allow mining and/or downloading of raw data.
8	☐	RAW DATA CATALOG: The format of the dataset is one of the following: XML, CSV/TXT, KML/KMZ, Excel (XLS), ESRI Shapefile or in another machine readable format. (Data in HTML and PDF files should NOT be considered for the "Raw" Data Catalog.)
9	☐	TOOLS CATALOG: If single or multiple raw datasets are offered within a "Tool" environment, the tool that offers the raw dataset(s) is one of the following: (1) Data Extraction Tool or webpage with downloadable datasets; (2) Feeds such as RSS, Atom or CAP; (3) a Widget (anything that requires a login or restricts use of data is prohibited.)
10	☐	The data owner and/or program office understand they are responsible for hosting data submissions and that they should provide an active URL which DATA.gov will only reference (i.e., no data is uploaded directly to DATA.gov.) *Please provide proposed Internet hosting location:*
11	☐	The data owner and/or program office agree to maintain the dataset and respond to all public comments.
12	☐	The data owner and/or program office will submit updates to the dataset, metadata, and necessary URL(s) in a timely manner.
13	☐	Complete and thoroughly describe the impact assessment of the proposed dataset submission (*Security, Privacy, FOIA, Legal)*:

** Authoritative data source is a recognized or official data production source with a designated mission statement or source/product to publish reliable and accurate data for subsequent users. An authoritative data source may be the functional combination of multiple or separate data sources.*

** High-value information is information that can be used to increase agency accountability and responsiveness; improve public knowledge of the agency and its operations; further the core mission of the agency; create economic opportunity; or respond to need and demand as identified through public consultation.*

Name: Requestor	**Office Symbol**	**Position**
Signature / Date	**Phone**	**E-mail**
Name: Data Owner (*Management Accountability*)	**Office Symbol**	**Position**

Signature / Date	**Phone**	**E-mail**
Name: OCIO Agency POC	**Office Symbol**	**Position**
Signature / Date	**Phone**	**E-mail**
Name: OCIO Executive Designee	**Office Symbol**	**Position**
Signature / Date	**Phone**	**E-mail**

8. Roles and responsibilities.

a. The program office is responsible for determining which datasets and tools are suitable to be posted on DATA.gov.

b. The program office retains the right and responsibility for managing its data and providing adequate technical documentation to include version control and archiving.

c. The program office is responsible for ensuring that data stewards for a particular dataset complete the required metadata.

d. The data steward is responsible for ensuring that the dataset is compliant with information and data quality guidelines in addition to completing an impact assessment.

e. The program office, in conjunction with the data steward, is responsible for ensuring that their datasets are consistent with statutory responsibilities including those related to security, accessibility, privacy, and confidentiality.

f. The program office and the CIO have the responsibility of ensuring that authoritative data sources are made available in formats that are platform independent and machine readable.

g. The CIO has the responsibility for assigning an overall DATA.gov POC for the agency.

h. The GSA POC is responsible for releasing datasets to DATA.gov along with facilitating comments back to program offices and data stewards.

9. DATA.gov submission process. The GSA DATA.gov submission process for candidate datasets to be published is described with regard to process scope, roles, decision criteria, and information flow. There are seven core process steps and three sub-processes that result from alternative decision paths. The process scope begins with any request for publication of a dataset and ends with the publication of the dataset on DATA.gov. The alternative decision path to end the process is to determine that GSA is not the authoritative source or the dataset is not suitable for publication. The following paragraphs coupled with Figure A-9 will illustrate each process step.

a. The GSA DATA.gov submission process begins when a customer requests publication of a dataset. A customer can be internal or external to GSA. If the request is from a Government agency, the DATA.gov PMO will forward the request to the GSA POC. The GSA POC determines if GSA is the authoritative source for the requested dataset. If it is, then the GSA POC identifies the appropriate data steward.

b. The data steward determines if the dataset is suitable for publication. Suitability is based on conformance to the information and data quality guidelines as well as GSA's DATA.gov checklist. If the dataset is suitable, the data steward obtains approval of the program office and completes the most current metadata template manually or through DMS.

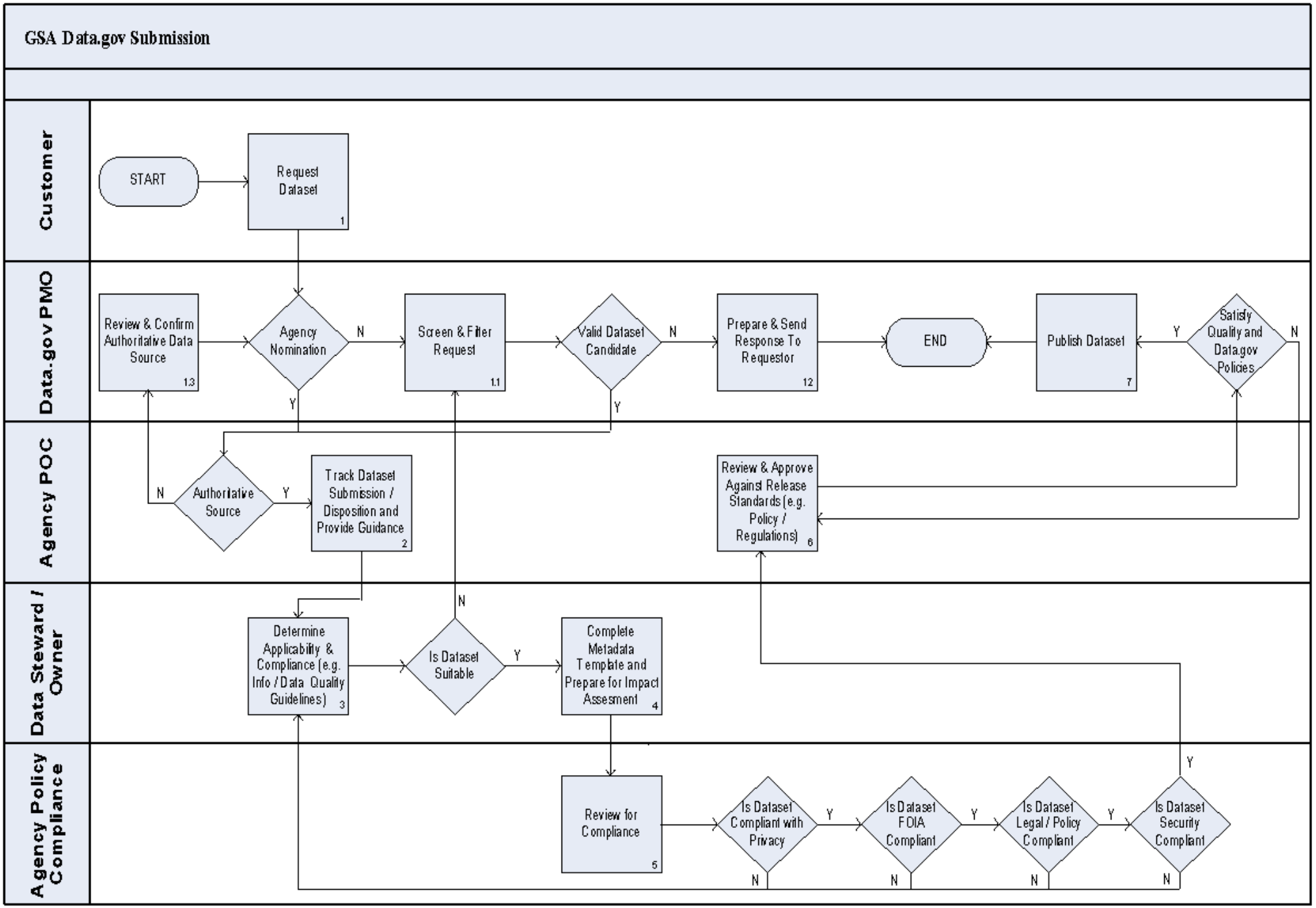
c. The data steward conducts an impact assessment on the dataset request package to test compliance with privacy, FOIA, legal, and security considerations. Upon a successful impact assessment, the dataset proposal is forwarded to the GSA POC.

d. The GSA POC then reviews the completed document package and forwards it to the DATA.gov PMO for publication to the DATA.gov web site. Upon publication the core process is completed.

e. The alternative process is engaged under three separate conditions that occur at subsequent points in the process:

- First, if the initial request is a non-agency nomination and does not pass the initial screening and filtering requirements, the request is rejected and the dataset is not published.
- Second, if the dataset request is valid and passes the initial screening but it is determined by the GSA POC that their agency is not the authoritative source for the dataset, the request would then be forwarded to the DATA.gov PMO to identify the correct authoritative source for the dataset.
- Third, if the dataset request is valid, and the agency is the authoritative source but the data steward deems the dataset not suitable based on the information and data quality guidelines and/or impact assessment, the request is rejected and the dataset is not published.

FIGURE A-9 GSA's DATA.gov Submission Process


GSA Data.gov Submission
Customer
Data.gov PMO
Agency POC
Data Steward / Owner
Agency Policy Compliance
START
Request Dataset
1
Review & Confirm Authoritative Data Source
1.3
Agency Nomination
N
Y
Screen & Filter Request
1.1
Valid Dataset Candidate
N
Y
Prepare & Send Response To Requestor
12
END
Publish Dataset
7
Y
Satisfy Quality and Data.gov Policies
N
Authoritative Source
N
Y
Track Dataset Submission / Disposition and Provide Guidance
2
Review & Approve Against Release Standards (e.g. Policy / Regulations)
6
Determine Applicability & Compliance (e.g. Info / Data Quality Guidelines)
3
Is Dataset Suitable
N
Y
Complete Metadata Template and Prepare for Impact Assesment
4
Review for Compliance
5
Is Dataset Compliant with Privacy
Y
N
Is Dataset FOIA Compliant
Y
N
Is Dataset Legal / Policy Compliant
Y
N
Is Dataset Security Compliant
Y
N

APPENDIX B:

Glossary of Relevant Terms

Term	Description
Authoritative Source	An authoritative data source is a recognized or official data production source with a designated mission statement or source/product to publish reliable and accurate data for subsequent users. An authoritative data source may be the functional combination of multiple or separate data sources.
Architecture	Representation of the structure of a system or community that describes the constituents of the system and how they interact with each other such that the goals and responsibility of the system or community are met.
Best Practice	A group of tasks that optimizes the efficiency or effectiveness of the business discipline or process to which it contributes. Best practices are generally adaptable and replicable across similar organizations or enterprises - and sometimes across different functions or industries.
Business Objective	Objectives state what is to be achieved, and the results and activities required to measure progress towards reaching the desired state.
Business Process	A business process is one aspect of a business model intended to specify the services, participants, interactions, resources and course of activities required to realize business value. A business process is a set of linked activities that create value by transforming an input into a more valuable output. Both input and output can be artifacts and/or information and the transformation can be performed by human actors, organizations, machines, or both. A business process can be decomposed into activities that may be atomic (E.G. “Delete file”) or utilize sub-processes (E.G. “Build Ship”), which contribute to achieving the goal of the super-process. The analysis of business processes typically includes the mapping of processes and sub-processes down to activity level. A business process may specify how processes are currently executed or may specify a future-sate process intended to improve business value and/or reduce costs.
Data	Data relates to a fact, event, or transaction. (1) Data itself has no meaning, but may become information when it is interpreted; and (2) Data can be symbols to include words (text and/or verbal), numbers, diagrams, and images (still &/or video).
Data Model	The main aim of a data model is to support the development of information systems by providing the definition and format of data.
Dataset	A dataset is an organized collection of data. The most basic representation of a dataset is data elements presented in tabular form. A dataset may also present information in a variety of non-tabular formats, such as an extended mark-up language (XML) file, a geospatial data file, or an image file, etc

Term	Description
Enterprise Architecture (EA)	A process and the associated strategic information asset base to support enterprise objectives, which includes: - the mission; - information necessary to perform the mission; - the technologies necessary to perform the mission; - the transitional processes; and - the business model for supporting the transformation of an enterprise to meet its changing business objectives.
Federal Enterprise Architecture (FEA) Data Reference Model (DRM)	The Data Reference Model (DRM) describes, at an aggregate level, the data and information supporting government program and business line operations. This model enables agencies to describe the types of interaction and exchanges occurring between the Federal government and citizens. The DRM categorizes government information into greater levels of detail. It also establishes a classification for Federal data and identifies duplicative data resources. A common data model will streamline information exchange processes within the Federal government and between government and external stakeholders. The DRM provides a standard means by which data may be described, categorized, and shared. These are reflected within each of the DRM's three standardization areas: - Data Description: Provides a means to uniformly describe data, thereby supporting its discovery and sharing - Data Context: Facilitates discovery of data through an approach to the categorization of data according to taxonomies; additionally, enables the definition of authoritative datasets within a community of interest (COI) - Data Sharing: Supports the access and exchange of data where access consists of ad-hoc requests (such as a query of a dataset), and exchange consists of fixed, re-occurring transactions between parties
Information	The term "information" means any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms. Information is data that has been processed in such a way as to be meaningful to the person who receives it.
Information Architecture	The framework for organizing the planning and implementation of information resources. The set of information, processes, and technologies that an enterprise has selected for the creation and operation of information systems.
Information Model	Shows the relationships or linkages between major areas of interest to the business creates a shared business vocabulary (e.g. semantic model), defining a community's agreement on important concepts and relationships between those concepts.
Information Life Cycle	The stages through which information passes, typically characterized as creation or collection, processing, dissemination, use storage, and disposition.

Term	Description
Information Technology (IT)	The term 'information technology', with respect to an executive agency means any equipment or interconnected system or subsystem of equipment, that is used in the automatic acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of data or information by the executive agency.
Integrity	In context of data and information it implies complete or whole structure. The most important aspect of data integrity per the data architecture discipline is to expose the data, the functions and the data's characteristics
Metadata	Semantic information associated with a given variable; includes business definitions of the data and clear, accurate descriptions of data types, potential values, original source system, data formats, and other characteristics. Metadata defines and describes business data. Examples of metadata include data element descriptions, data type descriptions, attribute/property descriptions, range/domain descriptions, and process/method descriptions according to the International Organization for Standardization (ISO) 11179-3.
Metric	A standard for measurement.
Objectivity	Involves a focus on ensuring that information is accurate, reliable and unbiased and that information products are presented in an accurate, clear, complete and unbiased manner
Reproducibility	Means that the information is capable of being substantially reproduced, subject to an acceptable degree of imprecision.
Service or Staff Office	A Program Office within GSA responsible for coordinating nationwide programs and supporting Federal agencies and citizen-oriented organizations.
Transparency	A quality or characteristic of data or information. Transparency Promotes accountability and provides information for citizens about what their Government is doing. Transparency can strengthen the connections between government agencies and the public they serve. Transparency helps ensure meaningful and informed public participation.
Utility	The usefulness of the information to the intended users. The data provider should stay informed of changing information needs and develop new data, models, and information products where appropriate.

APPENDIX C:

References

1) Federal Enterprise Architecture (FEA) - https://obamawhitehouse.archives.gov/omb/e-gov/FEA

2) GSA Enterprise Architecture Transition Strategy and Sequencing Plan http://ea.gsa.gov/

3) GSA Information Quality Guidelines-Section 515. https://www.gsa.gov/reference/reports/information-quality-guidelinessec-515/data-quality-guidelines

4) OMB Circular A-130 Revised - https://www.cio.gov/policies-and-priorities/circular-a-130/

5) DATA.gov Concept of Operation v0.93