



# MANUAL TRANSMITTAL

Department of the Treasury  
Internal Revenue Service

2.120.11

MAY 29, 2024

## EFFECTIVE DATE

(05-29-2024)

## PURPOSE

- (1) This transmittal revises IRM Part 2 Chapter 120 Engineering, Section 11 Data Management dated January 20, 2023 due to be OneSDLC Compliant.

## MATERIAL CHANGES

- (1) Updated IRM to be OneSDLC Compliant.

## EFFECT ON OTHER DOCUMENTS

IRM 2.120.11 dated January 20, 2023 is superseded.

## AUDIENCE

This process description is applicable to all projects following the One Solution Delivery Lifecycle (OneSDLC) and customers who request data management services through the Configuration Management, Change Management KISAM, the Work Request Management System (WRMS), or the OnLine 5081 Service.

Rajiv Uppal  
Chief, Information Officer



2.120.11

Data Management Process Description

## Table of Contents

### 2.120.11.1 Program Scope and Objectives

#### 2.120.11.1.1 Background

##### 2.120.11.1.1.1 Process Description

##### 2.120.11.1.1.2 Goal

##### 2.120.11.1.1.3 Objectives

#### 2.120.11.1.2 Authority

#### 2.120.11.1.3 Roles and Responsibilities

#### 2.120.11.1.4 Program Management and Review

#### 2.120.11.1.5 Program Control

##### 2.120.11.1.5.1 Controls

##### 2.120.11.1.5.2 Metrics

##### 2.120.11.1.5.3 Tailoring Guidelines

##### 2.120.11.1.5.4 Quality Assurance

#### 2.120.11.1.6 Terms/Definitions/Acronyms

##### 2.120.11.1.6.1 Terms and Definitions

##### 2.120.11.1.6.2 Acronyms

#### 2.120.11.1.7 Related Resources

#### 2.120.11.1.8 Training

### 2.120.11.2 Process Workflow

#### 2.120.11.2.1 Main Process Diagram

#### 2.120.11.2.2 Inputs

#### 2.120.11.2.3 Outputs

#### 2.120.11.2.4 Activities

#### 2.120.11.2.5 Procedure

##### 2.120.11.2.5.1 Step 1.0: Data Collection

###### 2.120.11.2.5.1.1 Cross-Functional Flow Diagram

##### 2.120.11.2.5.2 Step 2.0: Data Consolidation

###### 2.120.11.2.5.2.1 Cross-Functional Flow Diagram

##### 2.120.11.2.5.3 Step 3.0: Data Certification

###### 2.120.11.2.5.3.1 Cross-Functional Flow Diagram

##### 2.120.11.2.5.4 Step 4.0: Data Connection

###### 2.120.11.2.5.4.1 Cross-Functional Flow Diagram

##### 2.120.11.2.5.5 Step 5.0: Data Consumption

###### 2.120.11.2.5.5.1 Cross-Functional Flow Diagram



2.120.11.1  
(05-29-2024)  
**Program Scope and Objectives**

- (1) Overview - The objectives of this Data Management Process which is one of the major process of Data Engineering is to provide a set of related activities that accomplish a common goal of Data Management. The process definition laid out in this document further breaks down these Activities into Tasks, each of which have a complete set of attributes defined such as data and tool specifications and the role(s) responsible for executing the tasks. The document also includes process goal and objectives, metrics, role definitions, policies and other process related attributes.
- (2) Purpose - This IRM contains procedural steps for the Data Management Process within IRS.
- (3) Audience - This process description is applicable to all projects following the One Solution Delivery Lifecycle(OneSDLC) and customers who request data management services through the Change Management KISAM system, the Work Request Management System (WRMS), or the OnLine 5081 Service.
- (4) Policy Owner -The ACIO, Enterprise Services is responsible for overseeing all aspects of the Data Management Process .
- (5) Program Owner - The Chief of Engineer Process Maturity Group under the Director of Enterprise Services, Solution Engineering Division is responsible for the administration, procedures, and updates related to the Data Management IRM.
- (6) Primary Stakeholders - IT Organizations and Projects that have data management requirements
- (7) Program Goals - This IRM provides the fundamental knowledge and procedural guidance for employees who request the Data Management Services from Data Engineering Group. By following the processes and procedures of this IRM, employees will achieve the management of project related data.

2.120.11.1.1  
(01-20-2023)  
**Background**

- (1) Data Management process used to be one of the major processes of the Data Engineering chapter of the IRM. However, due to Solution Engineering IRM Restructure, the Data Management process becomes one of the processes within the Engineering chapter of the IRM. The Data Management process includes 5 core components which are Collect, Consolidate, Certify, Connect, and Consume. These 5C's are the cornerstone of Data Management.
  - Collect data that exist in a database (structured) and that exist outside the database (unstructured).
  - Consolidate and condense disjointed data/project into that which can be homogenized and that which can remain distinct so the relevant data will be easily accessible.
  - Certify the designated sources, processes, projects, while understanding the data quality, and governing the data assets.
  - Connect the data to each other as well as users to the data. Building out the metadata repository helps expose the data elements to the larger audiences.
  - Consume data by presenting the data in the proper/desired state for the recipient to efficiently make correct decisions. This effectively turns the data into information that then can be strategically used as knowledge.

- 2.120.11.1.1.1  
(01-20-2023)  
Process Description

(1)

The mission of Data Management is to define an enterprise-wide data environment to more easily and efficiently organize, identify, share, reuse, and correlate data that enables the business to consume information and maximize the value to the agency. This Data Management Process Description describes **what** happens within the process and provides an operational definition of the major components of the process. This description specifies in a complete, precise, and verifiable manner, the requirements, design, and behavior characteristics of Data Management process. The process description (PD) is a documented expression of a set of activities performed to achieve a given purpose.

Tailoring of this process in order to meet the individual needs of each project is covered in the Tailoring Guidelines section of this document. For the purpose of this document, roles are provided to describe a set of responsibilities for performing a particular set of related activities. Roles and/or responsibilities should fit your business terminology.
- 2.120.11.1.1.2  
(01-20-2023)  
Goal

(1)

The process goal describes a specific purpose or achievement toward which the efforts of the process are directed. Each process has a specific focus and when combined with the other processes, forms a comprehensive framework for delivering and managing services.

This IRM provides the fundamental knowledge and procedural guidance for employees who request Data Management Services from the Data Engineering Group.

  - To achieve a given purpose and states the guidelines that all projects should follow regarding the Data Management process.
- 2.120.11.1.1.3  
(01-20-2023)  
Objectives

(1)

Process objectives describe material outcomes that are produced or achieved by the process. The following is a list of objectives for this process:

  - Provide standardized and institutionalized Data Management guidelines within IRS.
- 2.120.11.1.2  
(01-20-2023)  
Authority

(1)

All proposed changes to this document must be submitted in writing, with supporting rationale, to the Chief, Enterprise Service, Solution Engineering, Engineer Process Maturity Group.
- 2.120.11.1.3  
(01-20-2023)  
Roles and Responsibilities

(1)

Each process defines at least one role. Each role is assigned to perform specific tasks within the process. The responsibilities of a role are confined to the specific process. They do not imply any functional standing within the hierarchy of an organization. For example, the process manager role does not imply the role is associated with or fulfilled by someone with functional management responsibilities within the organization. Within a specific process, there can be more than one individual associated with a specific role. Additionally, a single individual can assume more than one role within the process although typically not at the same time.

The following roles have been identified for this process:

Name	Description
------	-------------

Project Office	<ul style="list-style-type: none"> <li>Creates and submits Change Request or Work Request</li> <li>Enters OneSDLC process</li> <li>Conducts data profiling; implements data quality check; develops balance control and validation</li> <li>Submits an OnLine 5081 request for access the data system</li> <li>Submits a Live Data Waiver request as required user for SBU data</li> <li>Consumes data through queries and generate required reports</li> </ul>
High Authority(Infrastructure Executive Steering Committee(I-ESC))	Provides governance for data technology platforms and funds data initiatives
Engineer/Analyst (Data Engineering Group)	<ul style="list-style-type: none"> <li>Determines scope and impact as related to ELDM</li> <li>Determines data organization (OLTP/OLAP)</li> <li>Advances OneSDLC and achieves milestone approval</li> <li>Reviews Data Management Process outputs</li> <li>Aligns Project Data Model with ELDM</li> <li>Determines impact on Data Warehouse environment</li> <li>Certifies data contents</li> <li>Identifies interfaces</li> <li>Routes approved CRs to data platforms (e.g. EIP and BOE)</li> </ul>
OnLine 5081 Approval Group	<ul style="list-style-type: none"> <li>5081 Application/Services Administrator approves 5081 services request</li> </ul>

2.120.11.1.4  
(01-20-2023)  
**Program Management  
and Review**

- (1) Policies outline a set of plans or courses of action that are intended to influence and determine decisions or actions of a process. Policies provide an element of governance over the process that provides alignment to business vision, mission and goals.

<b>Process Management</b>	
---------------------------	--

Statement:	The Data Management process will have a single Process Owner and a separate Process Manager responsible for implementation and ensuring adherence to the process. The process will be reviewed regularly to ensure that it continues to support the business requirements of the enterprise. The process will be designed and developed based on ROI to the business. Process metrics will be focused on providing relevant information as opposed to merely presenting raw data.
<b>People:</b>	
Statement:	Roles and responsibilities for the process must be clearly defined and appropriately staffed with people having the required skills and training. The mission, goals, scope and importance of the process must be clearly and regularly communicated by upper management to the staff and business customers of IT. All IT staff (direct and indirect users of the process) shall be trained at the appropriate level to enable them to support the process.
Rationale:	It is imperative that people working in, supporting or interacting with the process in any manner understand what they are supposed to do. Without that understanding the Data Management process will not be successful.
<b>Process:</b>	

Statement:	Modifications to the Data Management process must be approved by the Process Owner. The design of the process must include appropriate interfaces with other processes to facilitate data sharing, escalation and workflow. The process must be capable of providing data to support real-time requirements as well as historical/ trending data for overall process improvement initiatives. The process must be fully documented, published and accessible to the various stakeholders of the process. The process will be reviewed on a periodic basis in order to ensure it continues to support organizational goals and objectives (continuous improvement). The process must include Inputs, Outputs, Controls, Metrics, Activities, Tasks, Roles and Responsibilities, Tool and Data requirements along with documented process flows. The process will be kept straight forward, rational, and easy to understand.
Rationale:	The process must meet operational and business requirements.
<b>Technology and Tools:</b>	
Statement:	<p>All tools selected must conform to the enterprise architectural standards and direction. Existing in-house tools and technology will be used wherever possible, new tools will only be entertained if they satisfy a business need that cannot be met by current in-house tools. The selection of supporting tools must be process driven and based on the requirements of the business. Selected tools must provide ease of deployment, customization and use. Automated workflow, notification and escalation will be deployed wherever possible to minimize delays, ensure consistency, reduce manual intervention and ensure appropriate parties are made aware of issues requiring their attention.</p> <p>The tools used by this process are the following::</p> <ul style="list-style-type: none"> <li>• BusinessObejcts</li> <li>• Enterprise Informatica Platform</li> <li>• Greenplum Database</li> <li>• Tableau</li> </ul>

Rationale:	Technology and tools should be used to augment the process capabilities, not become an end themselves.
------------	--

2.120.11.1.5  
(01-20-2023)  
Program Control

(1) Activities involved in ensuring a process is predictable, stable, and consistently operating at the target level of performance.

2.120.11.1.5.1  
(01-20-2023)  
Controls

(1) Process controls represent the policies and guiding principles on how the process will operate. Controls provide direction over the operation of processes and define constraints or boundaries within which the process must operate.

The process controls are the following:

Name	Description
Data Implementation Guide (DIG)	Data implementation begins with vision and strategy activities and ends with system deployment. The Data Implementation Guide (DIG) provides guidance on how to document the flow of an information system's data throughout its life cycle: from acquiring, conversion, migration, storage to the time when it becomes obsolete and is deleted.
Data Implementation Compliance Directive	The Data Implementation Compliance Directive enforces the Implementing of the Data Implementation Guide (DIG).
Scope	All projects following the One Solution Delivery Lifecycle(OneSDLC) are required to perform engineering processes and associated activities in accordance with the Engineering Policy IRM.
Change Management Policy	Change Management Policy is applicable to any change that might affect the IT systems, infrastructure, and services in the IT environment. This should also include changes to all architectures, applications, software, tools, and documentation, as well as changes to all configuration items across the whole service lifecycle.

2.120.11.1.5.2  
(01-20-2023)

**Metrics**

- (1) Metrics are used for the quantitative and periodic assessment of a process. They should be associated with targets that are set based on specific business objectives. Metrics provide information related to the goals and objectives of a process and are used to take corrective action when desired results are not being achieved and can be used to drive continual improvement of process effectiveness and efficiency.
- (2) Management will regularly set targets for process performance, gather quantifiable data related to different functions of the Data Management process, and review that data in order to make informed decisions and take appropriate corrective action, if necessary. All Measurements will have a defined data dictionary, map to the organizational strategic goals, and be documented in a Process Measurement Plan. The Process Measurement Plan template is available in the IT PAL.

2.120.11.1.5.3  
(01-20-2023)

**Tailoring Guidelines**

- (1) The tailoring guidelines identify the allowable variations of the IT organization's standard process as needed for adjustments (adding, deleting, modifying) relative to specific operational or functional needs of another organization. Process tailoring is about roles and procedures, not the standard process or major activities defined in this process. All tailoring request, with supporting rationale, must be submitted in writing to and approved by the Data Management Process owner.

2.120.11.1.5.4  
(01-20-2023)

**Quality Assurance**

- (1) The Data Management Process may be part of a Quality Assurance review at the discretion of the Process Owner.

2.120.11.1.6  
(01-20-2023)

**Terms/Definitions/  
Acronyms**

- (1) Terms/Definitions/Acronyms

2.120.11.1.6.1  
(01-20-2023)

**Terms and Definitions**

- (1) Terms and Definitions

Term	Definition
E300	The E300 is a subset of the E53, Capital Asset Plan, that summarizes the Business Case and outlines the benefits, costs, risks, and other OMB/Treasury requested data. It demonstrates an alignment with organizational objectives. It is used to determine if the effort should receive funding. The IRS submits the E300 to OMB to obtain funding approval for new major and existing major projects (Please see <b>IRM 2.16.1.3.4.2</b> ). The E300 is updated during the life of the project based on the annual budget cycle.

E53	The E53 is the Agency IT Investment Portfolio. It is submitted to request funds from OMB as part of the normal budgeting cycle and updated as required. All investments requiring OMB IT funding are listed on the E53.
Enterprise Logical Data Model (ELDM)	Enterprise Logical Data Model refers to an entity/relationship model in third normal form with fully attribute entities. It is a fully attributed entity relationship diagram (ERD), which shows each entity, its relationship to other entities and specifies the applicable business rules. An ELDM should focus only on business data. The value of logical data modeling has always been to establish the “single version of the truth” of an organization’s business functions.
KISAM	Knowledge Incident/Problem Service Asset Management System <a href="https://selfservice.web.irs.gov/webtier-9.52/ess.do">https://selfservice.web.irs.gov/webtier-9.52/ess.do</a>
RACI	The RACI model is based on the principle that people act in one of four ways when executing a task. It accounts for the fact that more than one role may be active in performing a specific task while clearly defining specific responsibilities for that role. While many roles may be involved in a task only one is Accountable for the results. The actions are: R Responsible for the action (may do the task) A Accountable for the action (including approval) C Required to be Consulted on the action I Required to be Informed of the action If a task does not have an Accountable role indicated then the Responsible role is assumed to be accountable for the task.
WRMS	The Work Request Management System is an ITIL-based COTS product for registering and managing demand requests for IT products and services.

2.120.11.1.6.2  
(01-20-2023)  
**Acronyms**

(1) Acronyms

Acronyms	Description
OneSDLC	One Solution Delivery Lifecycle
IPM	Integrated Process Management or Integrated Production Model when uses for database environment
ITIL	Information Technology Infrastructure Library
PAL	Process Asset Library
RACI	Responsible, Accountable, Consulted and Informed
RDBMS	Relational Database Management System
SBU	Sensitive But Unclassified

2.120.11.1.7  
(01-20-2023)

#### Related Resources

- (1) OneSDLC Artifacts.

2.120.11.1.8  
(01-20-2023)

#### Training

- (1) Process training involves training all stakeholders about key processes that are crucial for an organization to deliver business objectives. Training provides clarity to employees on a set of procedures that needs to be carried out as part of the process and the best possible way to do them. List below the training resources available for this process:  
Any Data Tools that come into the process.

- BusinessObjects
- Enterprise Informatica Platform
- Greenplum Database
- Tableau

2.120.11.2  
(01-20-2023)

#### Process Workflow

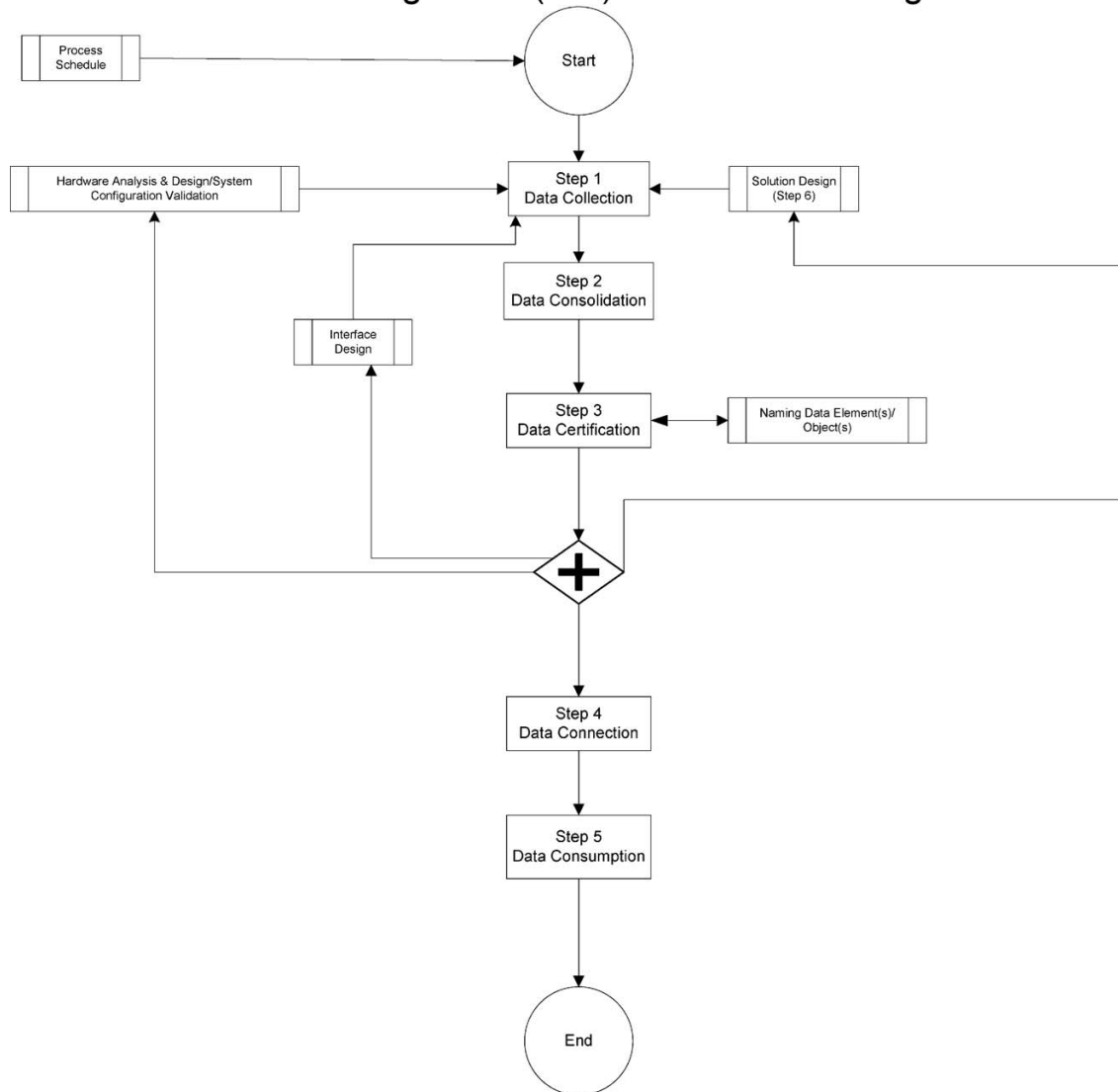
- (1) A process workflow consists of Activities and Tasks, Inputs and Outputs, Roles, and Flow Diagrams. It describes the tasks, procedural steps, organizations or people involved, required input and output information, and tools needed for each step of the process.

2.120.11.2.1  
(01-20-2023)

#### Main Process Diagram

- (1) Data Management Process Description Flow Diagram

### Data Management (DM) Process Flow Diagram



OS:IT:ES:SE:PM-PFD-DM Process Flow Diagram-V3.0-05212020

OS:IT:ES:SE:PM  
May 21, 2020**Figure 2.120.11-1**

2.120.11.2.2  
(01-20-2023)  
**Inputs**

- (1) Process inputs are used as triggers to initiate the process and to produce the desired outputs. Users, stakeholders or other processes provide inputs. The following is a list of inputs for this process:

Name	Description	Supplier
Change Request	A Change Request for the service of 5 C's from Data Management Process	Project Schedule (Project Office)
Disposition Change Request	Infrastructure Executive Steering Committee (IESC) would determine whether to approve or deny the request	Infrastructure Executive Steering Committee (IESC)
On-Line 5081 Request	A request for accessing the Data System	Project Office Schedule
Data Platforms	Solution Engineering Data Practice Group provides appropriate Data Platform to Project Office(s), such as: Enterprise Business Intelligence Platform (EBIP); Enterprise Informatica Platform (EIP)	Solution Engineering Data Engineering Practice Group (Engineer/Analyst)
Baseline and Milestone Updated for Primary Design	The following baseline documents should be started/created or updated with final signatures at the corresponding One Solution Delivery Life-cycle(OneSDLC) Phase & Milestone: <ul style="list-style-type: none"> <li>• E300 and E53</li> <li>• Project Tailoring Plan</li> <li>• Project Management Plan</li> <li>• Business System Report (BSR) (Logical Data Model/Business Metadata)</li> <li>• Agile Rapid Delivery Model (RDM) Vision, Scope, and Architecture (VSA), and all OneSDLC related artifacts</li> </ul>	Solution Engineering Data Engineering Practice Group (Engineer/Analyst)

Baseline and Milestone Updated for Detailed Design	<p>The following baseline documents should be started/created or updated with final signatures at the corresponding One Solution Delivery Life-cycle(OneSDLC) Phase &amp; Milestone:</p> <ul style="list-style-type: none"><li>• Design Specification Report (DSR) (Physical Data Model/Technical Metadata)</li><li>• Interface Control Document (ICD) (Data Movement Metadata)</li><li>• Simplified Design Specification Report (SDSR), and Agile all OneSDLC related artifacts</li></ul>	Solution Engineering Data Engineering Practice Group (Engineer/Analyst)
--	--	---

2.120.11.2.3  
(01-20-2023)  
Outputs

- (1) Each process produces tangible outputs. These outputs can take the form of products or data and can be delivered to a user or stakeholder or they can be used as inputs to other processes. Outputs are measurable in terms of quantity and quality.

Name	Description	Recipient
Data Platforms	Solution Engineering Data Practice Group provides appropriate Data Platform to Project Office(s), such as: Enterprise Business Intelligence Platform (EBIP); Enterprise Informatica Platform (EIP)	Project Office
Disposition Data Collection Change Request (CR)	Infrastructure Executive Steering Committee (IESC) determines if the Change Request is denied or approve	Project Office

Required Reports	Solution Engineering Data Practice Group provides all kinds of request reports from Project Office(s)	Project Office
Baseline and Milestone Updated for Primary Design	<p>A completed Data Management design with the applied version controlled:</p> <ul style="list-style-type: none"> <li>• E300 and E53</li> <li>• Project Tailoring Plan</li> <li>• Project Management Plan</li> <li>• Business System Requirement (BSR <ul style="list-style-type: none"> <li>a. Entities &amp; Attributes (Logical Data Model)</li> <li>b. Business Metadata (Business Taxonomy)</li> </ul> </li> </ul> <p>)</p> <ul style="list-style-type: none"> <li>• Agile Rapid Delivery Model (RDM) Vision, Scope, and Architecture (VSA),, and all OneSDLC related artifacts</li> </ul>	Project Office

Baseline and Milestone Updated for Detailed Design	<div>A completed Data Management design with the applied version controlled:<ul style="list-style-type: none"><li>• Design Specification Report (DSR) (Physical Data Model/ Technical Metadata)<ul style="list-style-type: none"><li>a. Tables &amp; Columns (Physical Data Model)</li><li>b. Technical Metadata (Technical Taxonomy)</li></ul></li><li>• Interface Control Document (ICD) (Data Movement Metadata)<ul style="list-style-type: none"><li>a. Schema; Service; Message Inter-change Metadata (XML Taxonomy)</li></ul></li><li>• Simplified Design Specification Report (SDSR), and all Agile OneSDLC related artifacts</li></ul></div>	Project Office
--	--	----------------

2.120.11.2.4  
(01-20-2023)  
**Activities**

- (1) An activity is a major unit of work to be completed in achieving the objectives of the process. A process consists of a sequence of related activities that transforms inputs into outputs and performed by the roles defined in the process. Activities are measurable in terms of efficiency and effectiveness. Identify the activities in the process and provide a brief description. The activities must correspond with the high-level process flow diagram above.

ID	Name	Description
Step 1.0	Data Collection	<p>To collect data that exist in databases (structured) and that exist outside databases (unstructured) from across the agency and external sources.</p> <ul style="list-style-type: none"><li>• Project Office creates and submits Data Collection Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the data source, and evaluates the Change Request against the IRS data strategy.</li><li>• High Authority dispositions Data Collection Change Request (CR) if that required funding.</li><li>• Project Office follows OneSDLC Guidance and achieves OneSDLC Phase</li><li>• Project Office creates and submits OneSDLC artifacts including Data Collection processing outputs</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) reviews Data Collection processing outputs.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) aligns Project data model with Enterprise Logical Data Model (ELDM), and Standard Meta Model (Business Taxonomy and Technical Metadata).</li></ul>

Step 2.0	Data Consolidation	<p>To manage data as follows: consolidate and condense disjointed data that can/ should be homogenized and recognize which data is truly distinct so the relevant data is most easily identifiable and accessible.</p> <ul style="list-style-type: none"><li>• Project Office creates and submits Data Consolidation Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the scope of Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group (Engineering/Analyst) determines the data source, and evaluates the Change Request against the IRS data strategy.</li><li>• High Authority dispositions Data Consolidation Change Request (CR) if that required funding.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) considers Change Request (CR) and determines the impact on the Data Warehouse environment.</li></ul>
----------	--------------------	--

Step 3.0	Data Certification	<p>To certify that the designated sources, processes, and projects understand data quality, and govern the data assets.</p> <ul style="list-style-type: none"><li>• Project Office creates and submits Data Certification Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the scope of Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the data source, and evaluates the Change Request against the IRS data strategy.</li><li>• High Authority dispositions Data Consolidation Change Request (CR) if that required funding.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) certifies data contents with an appropriate data Competency Center.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) assigns the responsibility of source data validation to Project Office/User.</li><li>• Project Office conducts data profiling, implements data quality check, and develops balance &amp; control and validation.</li><li>• Interface with external process Naming Data Element(s)/Object(s)</li></ul>
----------	--------------------	--

Step 4.0	Data Connection	<p>To connect the data to each other as well as users to the data. Building out the metadata repository helps expose data element to the larger IRS audience</p> <ul style="list-style-type: none"><li>• Project Office creates and submits Data Connection Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the scope of Change Request (CR) or Work Request (WR).</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) determines the data source, and evaluates the Change Request against the IRS data strategy.</li><li>• High Authority dispositions Data Consolidation Change Request (CR) if that required funding.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) identifies interfaces.</li><li>• Solution Engineering Engineer/Analyst (Data Engineering Group) routes approved Change Request (CR) to appropriate data platform (e.g. Enterprise Informatica Platform (EIP) Integrated Services, BusinessObjects Enterprise (BOE)).</li></ul>
----------	-----------------	--

Step 5.0	Data Consumption	<p>To consume data by presenting the data in the proper/ desired state for the recipient to efficiently make the right decisions. This effectively turns data into information that then can be strategically used.</p> <ul style="list-style-type: none"> <li>• Project Office submits an OnLine 5081 request for accessing to the data system.</li> <li>• Project Office submits a Live Data Waiver request as required for users of SBU data.</li> <li>• OnLine 5081 Approval Group approves or denies the OnLine 5081 request.</li> <li>• Solution Engineering Engineer/Analyst (Data Engineering Group) provides data environment for Project Monitor Control creates and submits Data to consume data.</li> <li>• Project Office creates and submits Data consumes data through queries and generating reports by using various data retrieval resources for data analysis.</li> </ul>
----------	------------------	--

2.120.11.2.5  
(01-20-2023)

(1) Procedure

#### Procedure

2.120.11.2.5.1  
(01-20-2023)

#### Step 1.0: Data Collection

(1) **Data Collection:** To collect structured and unstructured data from across the agency and external to the agency. By leveraging the additional data captured with e-Filing and modernized residual paper processing, more data can be shared for matching and targeting the best case. Data can be collected in forms of documents, records, models, emails, processes, RDBMS, business rules, or transaction systems.

ID	Task Name and Description	Role	RACI	Duties
----	---------------------------	------	------	--------

<b>Step 1.1</b>	Request specific data services	Project Office	R	Project Office submits a Change request/Work request ticket for data services
<b>Step 1.2</b>	<p>Determine scope and impact of Change Request or Work Request</p> <p>This task is to determine the scope and impact of Change Request or Work Request with respect to IRS Data Collection as related to the Enterprise Logical Data Model (ELDM)</p>	Engineer/Analyst (Data Engineering Group)	R	Ensure change or work request follows the Data Engineering Guidelines
<b>Step 1.3</b>	<p>Determine the data source, and evaluate the Change Request or Work Request against the IRS data strategy</p> <ul style="list-style-type: none"> <li>• Determine the authoritative data source</li> <li>• Determine data organization (OLTP/OLAP)</li> <li>• Determine feasibility of Data Collection Change Request (CR)</li> <li>• Evaluate Data Collection Change Request (CR) against IRS Data Strategy</li> </ul> <p>The purpose of this process step is to determine the authoritative data source, data organization (OLTP/OLAP), feasibility of Data Collection, and evaluate Data Collection CR against IRS Data Strategy</p>	Engineer/Analyst (Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group) is responsible for evaluating the change request against the IRS data strategy

<b>Step 1.4</b>	<p>Disposition Change Request or Work Request requiring funding</p> <p>Infrastructure Executive Steering Committee (IESC) disposition change request and ensure project has funding to support change request. Infrastructure Executive Steering Committee (IESC) approves or rejects CA (Project) CR request. Solution Engineering Data Practice Group returns rejected Change Request or Work Request back to Change Initiator/Analyst (Project Office) for further modification, or determines data source, feasibility and enforce IRS Data Strategy</p>	<p>High Authority</p> <p>Engineer/ Analyst(Data Engineering Group)</p>	<p>A</p> <p>R</p>	<p>High Authority reviews the change request, and determines to reject or approve the request</p> <p>Forward change or work request to High Authority</p>
<b>Step 1.5</b>	<p>One Solution Delivery Lifecycle(OneSDLC)</p> <ul style="list-style-type: none"> <li>Follow OneSDLC Guidance</li> <li>Advance through OneSDLC and achieve Milestone Approval</li> </ul> <p>The purpose of this process step is Project Office enters OneSDLC to manage and implement business change through system initiatives.</p>	Project Office	R	Project Office starts (baseline) or updates at the corresponding OneSDLC) Phase & Milestone

2.120.11.2.5.1	Internal Revenue Manual	Cat. No. 74890K (05-29-2024)
----------------	-------------------------	------------------------------

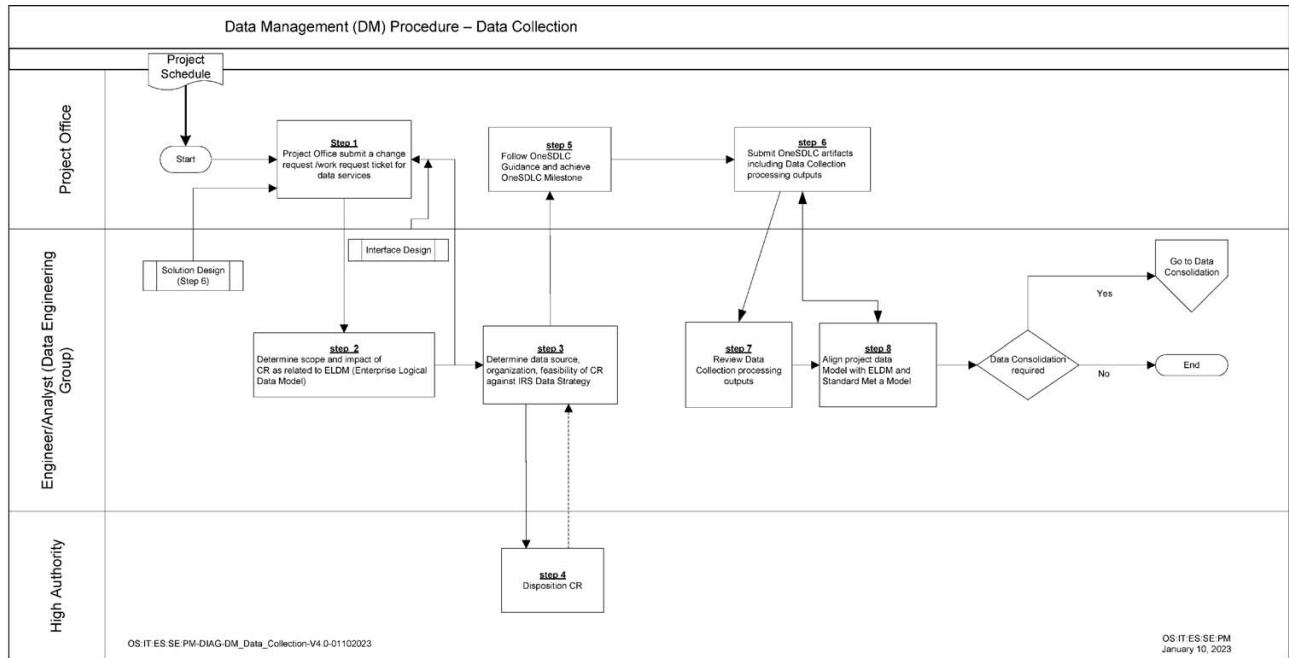
<b>Step 1.7</b>	<p>Review Data Collection Processing Outputs - Primary Design: E300 and E53 Project Tailoring Plan, Project Management Plan, Business System Requirement (BSR), or Agile Vision, Scope and Architecture (VSA), and all OneSDLC related artifacts.</p> <ul style="list-style-type: none"> <li>Entities &amp; Attributes (Logical Data Model)</li> </ul> <p>Business Metadata (Business Taxonomy) Detail Design: Design Specification Report (DSR) (Physical Data Model/Technical Metadata), or Agile Simplified Design Specification Report (SDSR), and all OneSDLC related artifacts.</p> <ul style="list-style-type: none"> <li>Tables &amp; Columns (Physical Data Model)</li> <li>Technical Metadata (Technical Taxonomy)</li> </ul> <p>Interface Control Document (ICD) (Data Movement Metadata)</p> <ul style="list-style-type: none"> <li>Schema; Service; Message Interchange Metadata (XML Taxonomy)</li> </ul> <p>The purpose of this process step is to ensure that OneSDLC artifacts including Data Collection Outputs for Solution Engineering review.</p>	<p>Engineer/Analyst(Data Engineering Group)</p>	<p>R</p>	<p>Review the OneSDLC artifacts which are submitted by Change Initiator/Change Analyst</p>
-----------------	--	---	----------	--

<b>Step 1.8</b>	<p>Align Project's data model with Enterprise Logical Data Model (ELDM), Standard Meta Model (Business Taxonomy and Technical metadata)</p> <ul style="list-style-type: none"> <li>• Collaborate with Change Initiator/Originator (Project Office)</li> <li>• Solution Engineering together with Project Office work to harmonize outputs until aligned with Enterprise Logical Data Model (ELDM)</li> <li>• Solution Engineering together with Project Office work to harmonize outputs until aligned with Standard Meta Model (Business Taxonomy and Technical metadata)</li> <li>• Determine if this Change Request or Work Request is only for this step or if there is a need to go to the next Data Consolidation step.</li> </ul> <p>The purpose of this process step is to ensure that the data model of Project Office aligned with Enterprise Logical Data Model (ELDM), Standard Meta Model (Taxonomy)</p>	<p>Engineer/ Analyst(Data Engineering Group)</p> <p>Project Office</p>	<p>R</p> <p>C</p>	<ul style="list-style-type: none"> <li>• Collaborates and works together with Project Office to harmonize the OneSDLC requested artifacts</li> <li>• Determines if Change Request/ Work Request only for Data Collection or would go to next Step - Data Consolidation</li> </ul> <p>Work with Solution Engineering Engineer/Analyst (Data Engineering Group) to harmonize OneSDLC artifacts until align with Enterprise Logical Data Model, and Meta Model.</p>
-----------------	---	--	-------------------	--

2.120.11.2.5.1.1

(1) Data Collection Cross-Functional Flow Diagram

(01-20-2023)

**Cross-Functional Flow Diagram****Figure 2.120.11-2**

2.120.11.2.5.2

(1) **Data Consolidations:** To manage data as follows: consolidate and condense disjoint data that can/should be homogenized and recognize which data is truly distinct to the relevant data is most easily identifiable and accessible.

(01-20-2023)

**Step 2.0: Data Consolidation**

ID	Task Name and Description	Role	RACI	Duties
<b>Step 2.1</b>	Request specific data services	Project Office	R	Project Office submits a Change/Work request ticket for data services

<b>Step 2.2</b>	<p>Determine the scope of Change Request or Work Request</p> <ul style="list-style-type: none"> <li>• Before processing this step determine if Data Collection step needed, if yes go to Data Collection Steps; otherwise;</li> <li>• Determine the scope and impact of Change Request or Work Request with the respect to IRS Data Consolidation as related to the Enterprise Logical Data Model (ELDM)</li> </ul>	Engineer/ Analyst(Data Engineering Group)	R	Ensure change or work request follows the Data Engineering Guidelines
<b>Step 2.3</b>	<p>Determine the data source, and evaluate the Change Request against the IRS data strategy</p> <ul style="list-style-type: none"> <li>• Determine the authoritative data source</li> <li>• Determine data organization (OLTP/OLAP)</li> <li>• Determine feasibility of Data Collection Change Request (CR)</li> <li>• Evaluate Data Collection Change Request (CR) against IRS Data Strategy</li> </ul>	Engineer/ Analyst(Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group) is responsible for evaluating the change request against the IRS data strategy

<b>Step 2.4</b>	Disposition change request requiring funding.  Infrastructure Executive Steering Committee (IESC) disposition change request and ensure project has funding to support change request	High Authority  Engineer/Analyst(Data Engineering Group)	A  R	High Authority reviews the change request, and determines to reject or approve the request       Forward change or work request to High Authority
<b>Step 2.5</b>	Consider the Data Consolidation Change Request or Work Request, and determine the impact on the data warehouse environment <ul style="list-style-type: none"> <li>• Determine the impact on the data warehouse environment</li> <li>• Determine if this Change Request or Word Request is only for this step or if there is a need to go to the next Data Certification step</li> </ul>	Engineer/Analyst (Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group would consider the Change Request, and determine the impact on the data warehouse environment

2.120.11.2.5.2.1  
(01-20-2023)

(1) Data Consolidation Cross-Functional Flow Diagram

#### Cross-Functional Flow Diagram

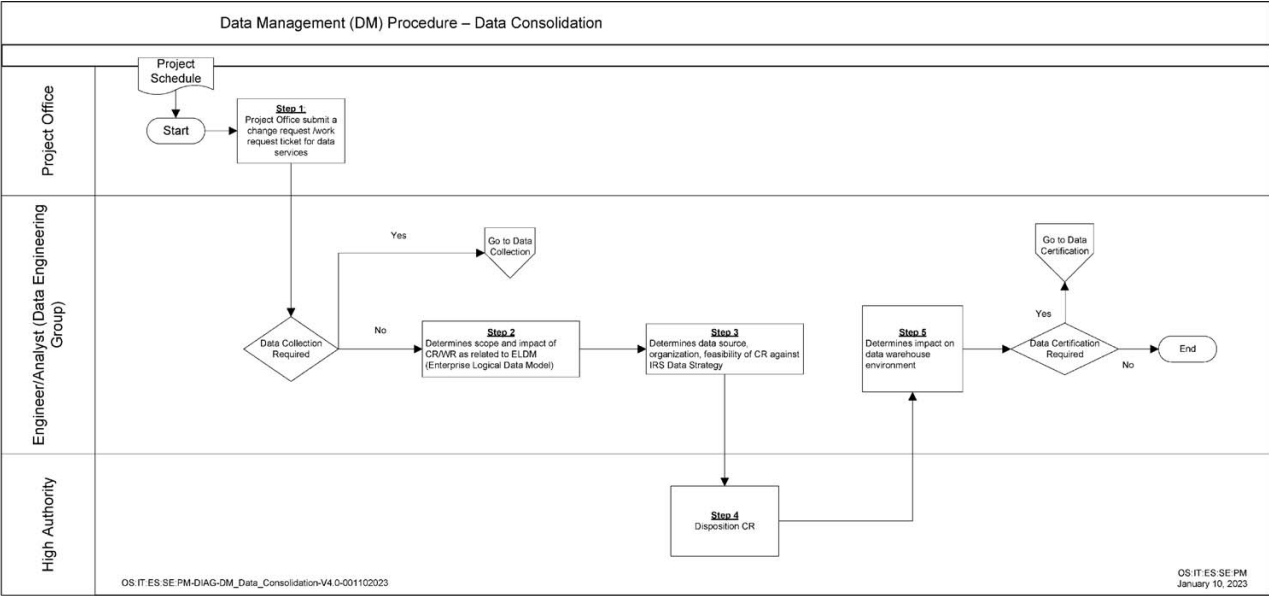


Figure 2.120.11-3

2.120.11.2.5.3  
(01-20-2023)  
**Step 3.0: Data  
Certification**

- (1) **Data Certification:** To certifies the designated sources, processes, projects and understand the data quality. The higher the quality of the data the Agency has, the more accurate scores can be produced. More accurate scores lead to more effective issue detection and case selection. Certification enables the governance of complete, consistent, comprehensible, correct, and authoritative sources by establishing operational metrics.

ID	Task Name and Description	Role	RACI	Duties
Step 3.1	Request specific data services	Project Office	R	Project Office submits a Change/ Work request ticket for data services

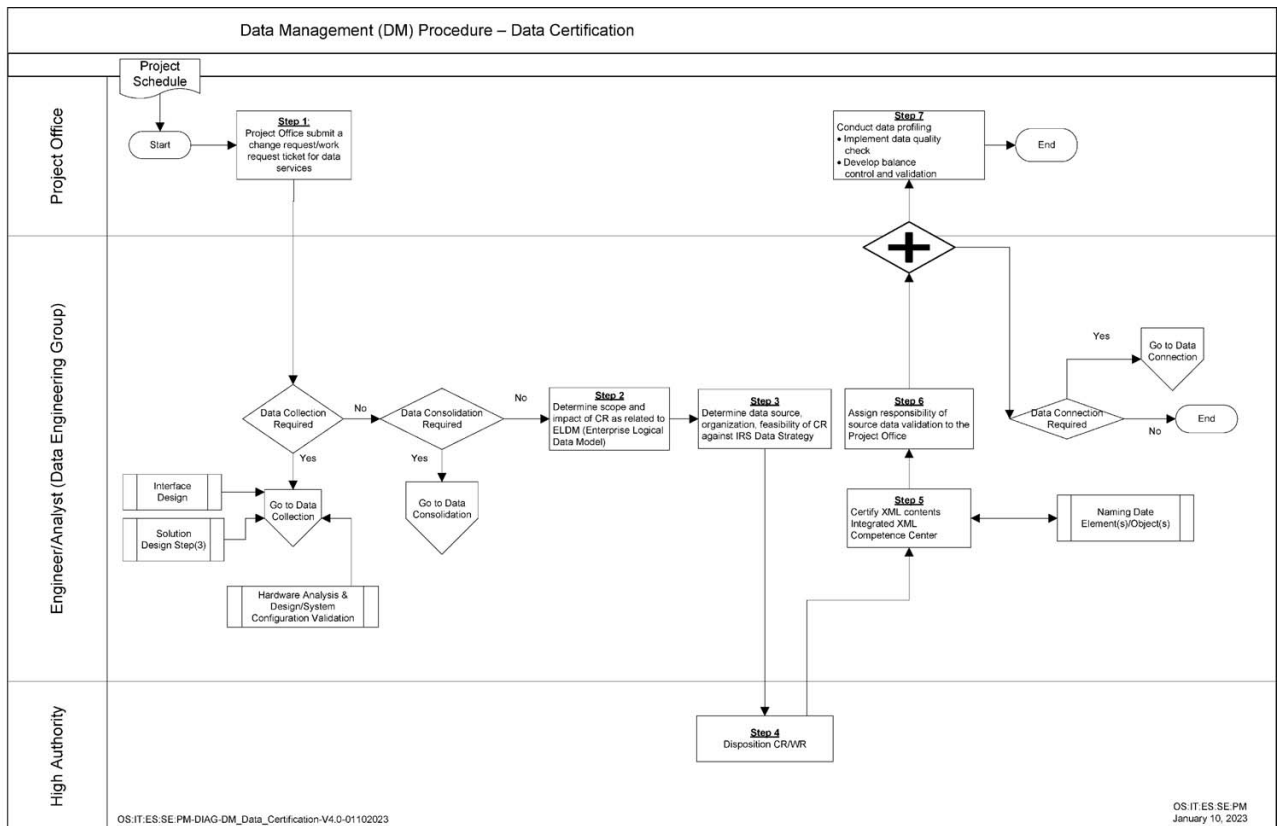
<b>Step 3.2</b>	<p>Determine the scope of Change Request or Work Request</p> <ul style="list-style-type: none"> <li>Before processing this step, determine if other Data Management activities such as: Data Collection step or/and Data Consolidation step are needed. If Data Collection needed, interfaces three external processes, such as: Interface Design, Solution Design (step 6), and Hardware Analysis &amp; Design/ System Configuration Validation</li> <li>If Data Consolidation step needed, go to Data Consolidation Step; otherwise;</li> <li>Determine the scope and impact of Change Request or Work Request with the respect to IRS Data Certification as related to the Enterprise Logical Data Model (ELDM)</li> </ul>	Engineer/ Analyst(Data Engineering Group)	R	Ensure change or work request follows the Data Engineering Guidelines
<b>Step 3.3</b>	<p>Determine the data source, and evaluate the Change Request against the IRS data strategy</p> <ul style="list-style-type: none"> <li>Determine the authoritative data source</li> <li>Determine data organization (OLTP/OLAP)</li> <li>Determine feasibility of Data Certification Change Request (CR)</li> <li>Evaluate Data Certification Change Request (CR) against IRS Data Strategy</li> </ul>	Engineer/ Analyst(Data Engineering Group)	R	Solution Engineering Engineer/ Analyst (Data Engineering Group) is responsible for evaluating the change request against the IRS data strategy

<b>Step 3.4</b>	Disposition change request requiring funding.	High Authority	A	High Authority
	Infrastructure Executive Steering Committee (IESC) disposition Change Request or Work Request and ensure project has funding to support change request	Engineer/ Analyst(Data Engineering Group)	R	Forward change or work request to High Authority
<b>Step 3.5</b>	Interfaces with external Naming Data Element(s)/ Object(s) to certify the data content with appropriate data Competency Center Based on data content, Solution Engineering Data Practice Group(Engineer/ Analyst) would send to different Competency Center to validate	Engineer/ Analyst(Data Engineering Group)	R	Solution Engineering Engineer/ Analyst (Data Engineering Group)) would certify data content to with competency center
<b>Step 3.6</b>	Assign the responsibility of source data validation to the Project Office <ul style="list-style-type: none"> <li>send data validation back to Project Office</li> <li>determine if this Change Request is only for this step or go to the next Data Management step(s)</li> </ul>	Engineer/ Analyst(Data Engineering Group)	R	After certifying data content, Solution Engineering Engineer/ Analyst (Data Practice Group) would assign the responsibility to project office
		Project Office	I	Change Initiator/ Change Analyst accepts the feedback from Solution Engineering Data Practice Group (Engineer/ Analyst) that the data have been validated
<b>Step 3.7</b>	Execute the following data validation: <ul style="list-style-type: none"> <li>Conduct data profiling</li> <li>Implement data quality check</li> <li>Develop balance &amp; control and validation</li> </ul>	Project Office	R	After data validation, Change Initiator/Originator would do validation and quality check

2.120.11.2.5.3.1  
(01-20-2023)

(1) Data Certification Cross-Functional Flow Diagram

**Cross-Functional Flow  
Diagram**



**Figure 2.120.11-4**

2.120.11.2.5.4  
(01-20-2023)

**Step 4.0: Data  
Connection**

- (1) **Data Connection:** To connect the data to each other as well as users to the data. Building out the metadata repository helps expose data elements to a larger audience. Increasing data accessibility across the agency helps to gain the ability to measure ROI and compliance impact across the end to end process.

ID	Task Name and Description	Role	RACI	Duties
Step 4.1	Request specific data services	Project Office	R	Project Office submits a Change/work request for data services

<b>Step 4.2</b>	<p>Determine the scope of Change Request or Work Request</p> <ul style="list-style-type: none"> <li>Before processing this step, determine if other Data Management activities such as: Data Collection step or/and Data Consolidation step or/and Data Certification step are needed If Data Collection needed, interfaces three external processes, such as: Interface Design, Solution Design (step 6), and Hardware Analysis &amp; Design/ System Configuration Validation If Data Consolidation step needed, go to Data Consolidation Step</li> <li>If Data Certification step needed, go to Data Certification Step; otherwise,</li> <li>Determine the scope and impact of Change Request or Work Request with the respect to IRS Data Consolidation as related to the Enterprise Logical Data Model (ELDM) .</li> </ul> <p>This task is to determine the scope and impact of Change Request with respect to IRS Data Connection as related to the Enterprise Logical Data Model (ELDM)</p>	Engineer/Analyst (Data Engineering Group)	R	Ensure change or work request follows the Data Engineering Guidelines
-----------------	---	---	---	---

<b>Step 4.3</b>	Disposition change request requiring funding.	High Authority	A	High Authority reviews the change request, and determine to reject or approve the request
	Infrastructure Executive Steering Committee (IESC) disposition change request and ensure project has funding to support change request	Engineer/Analyst (Data Engineering Group)	R	Forward change or work request to High Authority for decision
<b>Step 4.4</b>	Identify Interface	Engineer/Analyst (Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group) identifies Interface
<b>Step 4.5</b>	Route approved Change Request or Work Request to appropriate data platforms (e.g. Enterprise Informatica) <ul style="list-style-type: none"> <li>• Provide Computer Operator Handbooks (COH) for data platforms</li> <li>• Evaluate schema certification metrics using Standard Data Names (Business, Physical &amp; XML Platform)</li> <li>• Determine if this Change Request or Work Request is only for this step or go to the next step</li> </ul>	Engineer/Analyst (Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group) routes approved Change Request to appropriate data platforms

2.120.11.2.5.4.1  
(01-20-2023)

(1) Data Connection Cross-Functional Flow Diagram

#### Cross-Functional Flow Diagram

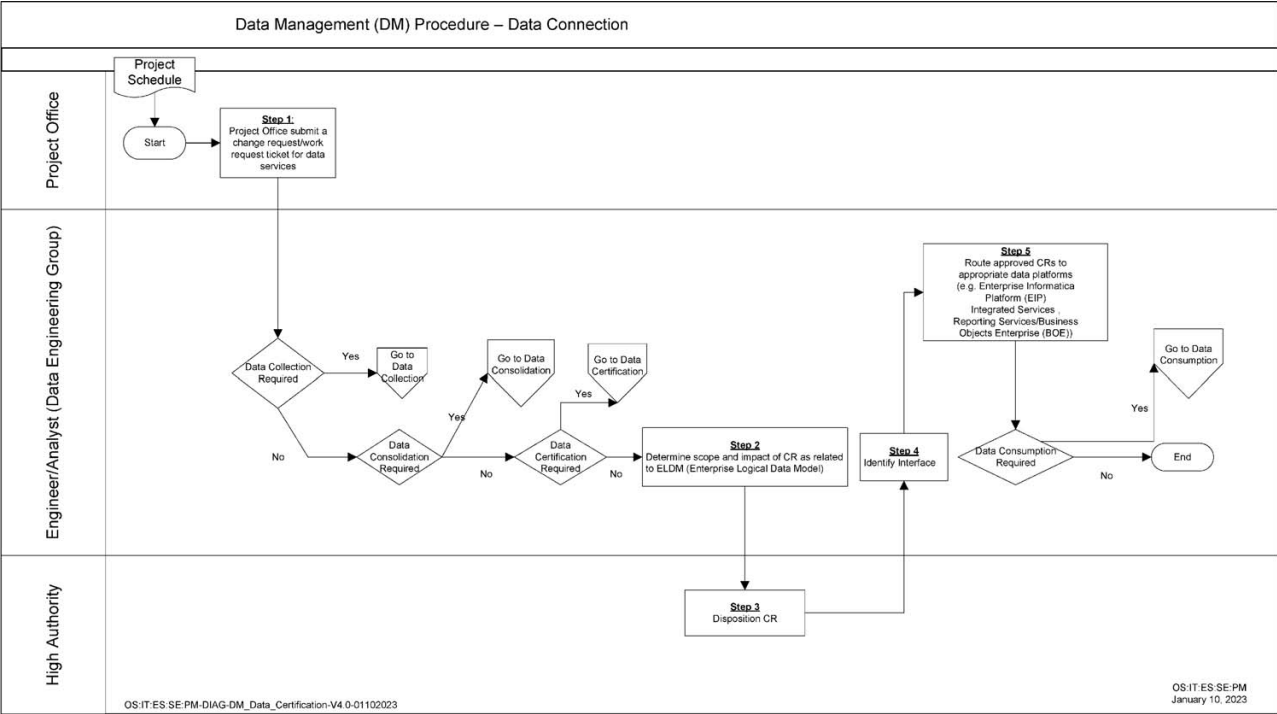


Figure 2.120.11-5

2.120.11.2.5.5  
(01-20-2023)  
**Step 5.0: Data  
Consumption**

(1) **Data Consumption:** To present the data in the proper/desired state for the recipient to efficiently make the right decision. Data is effectively turned into information, used to innovate new methods/discoveries, streamline and automate retrieval processes through renovations, and intent individuals through motivation. Tools of data consumption focus on visual representations of dashboards, scorecards, queries, reporting tools, OLAP, or data mining.

ID	Task Name and De- scription	Role	RACI	Duties
Step 5.1	Submit an OnLine 5081 request for access to the data system	Project Office	R	Project Office submits an OnLine 5081 request for access to the data system

<b>Step 5.2</b>	Submit a Live Data Waiver request as required for users of SBU data	Project Office	R	Project Office submits a Live Data Waiver request as required users for SBU data
<b>Step 5.3</b>	Approve or deny the OnLine 5081 request. If deny, end step; otherwise, if Data Collection required, go to Data Collection ; otherwise, if Data Consolidation required, go to Data Consolidation; otherwise, if Data Certification required, go to Data Certification; otherwise if Data Connection Required, go to Data Connection	OnLine 5081 Approval Group	R	OnLine 5081 System Administrator or Manager dispositions OnLine 5081 request
<b>Step 5.4</b>	Provide data environment for Project/User to consume data	Engineer/Analyst (Data Engineering Group)	R	Solution Engineering Engineer/Analyst (Data Engineering Group) provides data consumption environment
<b>Step 5.5</b>	Consume data through queries and generating reports by using various data retrieval resources for data analytic	Project Office	R	Consume data through queries

2.120.11.2.5.5.1  
(01-20-2023)

(1) Data Consumption Cross-Functional Flow Diagram

#### Cross-Functional Flow Diagram

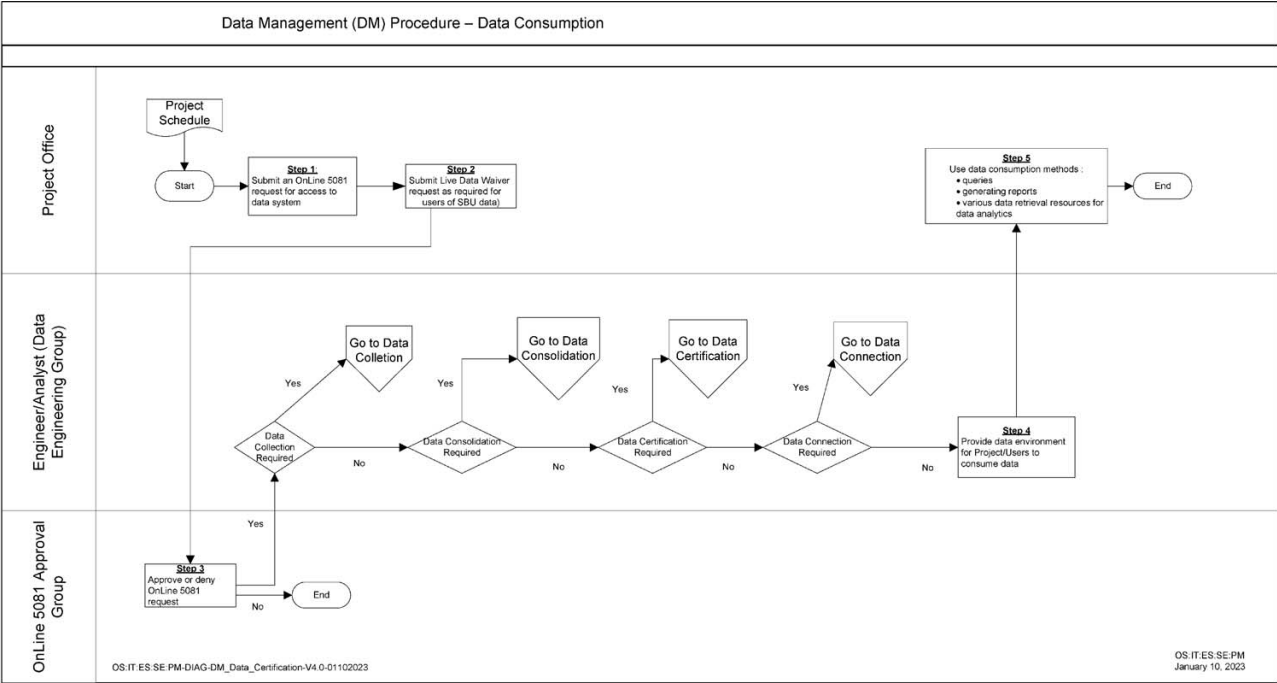


Figure 2.120.11-6