

Data Definitions Standards Guide

Carnegie Mellon Data Catalog
with Data Cookbook

Table of Contents





Overview.....	3
Roles and Responsibilities	3
Data Definition Approval Process	4
Data Definition Standards.....	5
1. Definition Name	6
2. Functional Definition.....	8
3. Other Information.....	11
4. Technical Definition	12
5. Tags.....	13
6. Functional Area.....	14

Overview

This guide is designed to provide standards for Data Definition Contributors as they create and submit new definitions for the university’s Data Catalog. Definitions are submitted through the [Data Cookbook](#), our central repository of university definitions.

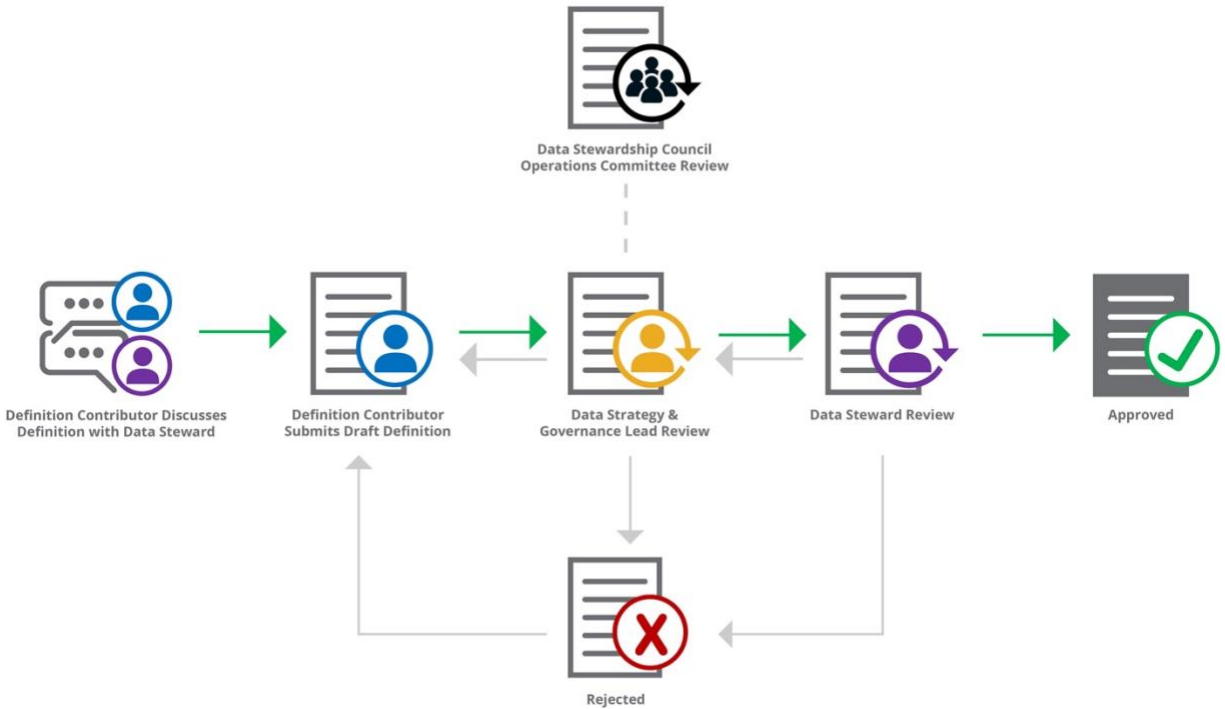
After Contributors submit a definition draft, the definition enters an approval workflow that includes review by the Data Strategy and Governance Lead and Data Stewards. The definitions are approved and then made available to the university community through the [Data Glossary](#).

Roles and Responsibilities

 Data Definition Contributor	The Data Definition Contributor is the subject matter expert for data within an enterprise administrative system. The Contributor initiates the definition approval workflow and is responsible for creating a definition draft and making any requested revisions.
 Data Strategy & Governance Lead	The Data Strategy & Governance Lead facilitates the approval process and reviews definitions to ensure they meet the Data Definition Standards. They serve as the liaison between the Data Definition Contributor and the Data Stewardship Council (DSC) Operations Committee as needed.
 Data Steward	The Data Stewards review the submitted data definition drafts for their specific Functional Areas and either approve, reject, or suggest edits.
 Data Stewardship Council (DSC) Operations Committee	The Data Stewardship Council Operations Committee is comprised of Data Stewards and representatives throughout the university who address issues related to conflicting institutional and administrative data definitions.

Data Definition Approval Process

When you submit a definition to the Data Cookbook, it enters an approval workflow. Each of the roles listed above has different responsibilities within the Approval workflow.



Data Definition Standards

Below is a completed definition in the Data Cookbook. There are six components to a definition:

1. Definition Name
2. Functional Definition
3. Other Information
4. Technical Definitions
5. Tags
6. Functional Areas

When you enter a definition into the Data Cookbook, you will need to adhere to the standards for each component, described in detail below.

The screenshot shows a web interface for a data definition. At the top, the title "Student Address Type" is marked with a red circle containing the number 1. Below the title, there is a green "Approved" badge and "Version 6". To the right, there are links for "Comment (0 comments)", "Add to my watch list", and "Show workflow".

The main content area is divided into sections:

- Functional Definition:** Marked with a red circle containing the number 2. It includes a "Business Definition" (A numeric code which identifies a particular type of address.), "Valid Values" (1 - Permanent Address, 2 - Campus Address, 3 - Mailing Address, 4 - Off Campus Address, 5 - Diploma Mailing Address), "Context and Cautions" (Permanent address is the student's permanent address while attending, and may not be a PO Box or academic department address. Campus address is the campus directory address for students while attending Carnegie Mellon. It includes assigned SMC Boxes for students who live in University Housing, and may also include on-campus addresses for graduate students. Off Campus address is the address where students live if they are not in University housing. Mailing Address is used by Undergraduate Admission and is not maintained after admitted students enroll. Diploma mailing address is an address collected only for the purposes of mailing a student's diploma. It is not maintained after that purpose.), "Links to Related Data Cookbook Definitions", "Data Steward" (Andrew Carnegie), "Data Sharing and Usage" (Used internally), and "Exception Criteria" (Any requests for additional use of this data should be sent to uro-sdw@andrew.cmu.edu.).
- Other Information:** Marked with a red circle containing the number 3. It includes "SOURCE OF DEFINITION: Data System" and "CLASSIFICATION: Private Data".
- Technical Definitions:** Marked with a red circle containing the number 4. It includes "Data System: SDW", "Time Context: Not Time Sensitive", and a note: "This technical definition requires additional permissions to view".

On the right side, there is a sidebar with several sections:

- FUNCTIONAL AREAS:** Marked with a red circle containing the number 6. It lists "University Registrar".
- WORKFLOW:** Marked with a red circle containing the number 6. It lists "Facilitated Definition Approval".
- MODERATORS:** Marked with a red circle containing the number 6. It lists "University Registrar Stewards".
- SYNONYMS:** Marked with a red circle containing the number 6. It lists "No Synonyms".
- TAGS:** Marked with a red circle containing the number 5. It lists "No Tags".
- ATTACHMENTS:** Marked with a red circle containing the number 6.
- RELATED SPECIFICATIONS:** Marked with a red circle containing the number 6. It lists "(0)".
- RELATED DEFINITIONS:** Marked with a red circle containing the number 6. It lists "(0)" and "No Related Definitions".
- COMMUNITY SHARING:** Marked with a red circle containing the number 6. It lists "(0)" and "There are no communities available. View your community settings here."

	Component	What is it?
1	Definition Name	The name given for the definition. When applicable, the associated Data System is listed in parenthesis.
2	Functional Definition	The business definition , along with the valid values, for the Data Element.
3	Other Information	The definition source and the data classification (i.e., public, private).
4	Technical Definitions	Where the Data Element can be located along with the data type and Data Element type . For example, this can include the database table and column name for the Data Element.
5	Tags	Searchable keywords and phrases that can be shared by multiple definitions .
6	Functional Areas	Topic or business area for the definition. Options include university entities such as Human Resources, University Advancement, the University Registrar, and others.

1. Definition Name

The Definition Name is a combination of the Data Element and, if applicable, the Data System, and uses the following format:

Data Element (Data System)

The Definition Name **must be unique** to the university's Data Cookbook. If your Definition Name already exists, review the definition to determine if your new definition is needed. You may also contact the Data Steward for the existing definition to discuss possible revisions or updates to the definition.

If your new definition is needed, include the Data System in the Name field following the guidelines below.

Create the Definition Name

2. In the Name field, enter the Data Element. The **Data Element** is the common name based on an item described in the report.

For example: The Data Element in a definition related to students' addresses might be Student Address Type.

3. (*Optional*) If applicable, enter the Data System in parenthesis after the Data Element. The **Data System** is the application or system where the Data Element exists. **Only include the Data System if the definition is generic or may originate in more than one data system** (e.g., Oracle, EBS, Workday, SPARCS).

For example: Student Address Type originates in more than one system. If the Student Address Type mentioned above is in the Student Data Warehouse (SDW), then the Data System would be SDW.

Tip: Select the same Data System when completing the Technical Definition section later in this process.

Example: Completed Definition Name

Using our example above, **if the Data Element Student Address Type exists in more than one application or system** our Definition Name would be:

Student Address Type (SDW)

Following the same example, **if the Data Element exists in only one application or system**, our Definition Name would be:

Student Address Type

Definition Name Standards

DO	DON'T
<ul style="list-style-type: none"> Make the name unique and descriptive so that it is distinguishable from other Definition Names. 	<ul style="list-style-type: none"> Do not begin the name with "A," "An," or "The."
<ul style="list-style-type: none"> Express the Definition Name in singular terms. 	<ul style="list-style-type: none"> Do not include numbers and special characters/symbols including underscores, commas, periods, and em-dashes.
<ul style="list-style-type: none"> Use common language that all audiences will understand. 	<ul style="list-style-type: none"> Do not enter extra spaces. Data Cookbook will create separate definitions if additional spaces are used. For example, "Account<space>Invoice," "Account<space space>Invoice," and "<space>Account Invoice" will be treated as different definitions.
<ul style="list-style-type: none"> Use Title Case by capitalizing the first letter of each word, with the exception of short words (e.g., a, the, by, for). 	

2. Functional Definition

The Functional Definition **describes and provides context for the Data Element**. It should explain the definition's uniqueness within its Functional Area and include links to related definitions, examples, and values. You should also provide context or perspective that would allow someone unfamiliar with the data to understand the definition.

Tip: When creating the Functional Definition, think about how it will be used and by whom. What do users need to know about it to use it correctly and effectively?

Create a Functional Definition

To create a Functional Definition, enter the following information into the Functional Definition field in the order listed below.

1. **Business Definition:** Write a concise explanation of the Data Element that uses common language and can be understood by someone unfamiliar with the element.
2. **Sample of Valid Values:** Describe how the Data Element is used and include hyperlinks to the authoritative source of valid values.
3. **Context and Cautions:** Include relevant additional information, such as context, cautions, common misinterpretations, historical changes in values or use, and sorting/grouping instructions.
4. **Links to Related Cookbook Definitions:** Link to relevant definitions that may provide context or related information for your definition. For example, the definition for ACT Composite Score includes links to definitions for the ACT Mathematics Score, ACT Natural Science Score, ACT Reading Score, etc.
5. **Keywords and Synonyms:** List terms that are commonly used in place of the Definition Name. (*Not shown in example below.*)
6. **Data Steward:** List the person who oversees data management and utilization for their enterprise administrative system. For a list of all Data Stewards, visit the [Data Stewardship Council](#).
7. **Data Sharing and Usage:** Include Information on how the data is used, both internally and externally, including reporting and integration information.
8. **Exception Criteria:** Provide information on how to request an exception to use the data outside of its accepted use. For instance, if a user wanted to include data from an internal source in a public-facing dashboard, the Exception Criteria would detail the process of requesting an exception from the Data Steward to use the data in this way.

Example: Completed Functional Definition

Below is a completed Functional Definition for Student Address Type that includes the information described above.

DEFINITION

Student Address Type

Approved Version 6

Functional Definition:

Business Definition 1
A numeric code which identifies a particular type of address.

Valid Values 2

- 1 - Permanent Address
- 2 - Campus Address
- 3 - Mailing Address
- 4 - Off Campus Address
- 5 - Diploma Mailing Address

Context and Cautions 3

Permanent address is the student's permanent address while attending, and may not be a PO Box or academic department address.

Campus address is the campus directory address for students while attending Carnegie Mellon. It includes assigned SMC Boxes for students who live in University Housing, and may also include on-campus addresses for graduate students.

Off Campus address is the address where students live if they are not in University housing.

Mailing Address is used by Undergraduate Admission and is not maintained after admitted students enroll.

Diploma mailing address is an address collected only for the purposes of mailing a student's diploma. It is not maintained after that purpose.

Links to Related Data Cookbook Definitions 4

Data Steward 5
Andrew Carnegie

Data Sharing and Usage 7
Used internally

Exception Criteria 8
Any requests for additional use of this data should be sent to uro-sdw@andrew.cmu.edu.

Other Information:

SOURCE OF DEFINITION: [Data System](#)

CLASSIFICATION: [Private Data](#)

[Comment \(0 comments\)](#)

[Add to my watch list](#)

[Show workflow](#)

FUNCTIONAL AREAS 1

University Registrar

WORKFLOW 1

Facilitated Definition Approval

MODERATORS 1

University Registrar Stewards

SYNONYMS 1

No Synonyms

TAGS 1

No Tags

ATTACHMENTS 1

RELATED SPECIFICATIONS (0) 1

RELATED DEFINITIONS (0) 1

No Related Definitions

COMMUNITY SHARING (0) 1

There are no communities available. View your community settings [here](#).

Functional Definition Standards

DO	DON'T
Use complete sentences and standard grammar .	Do not use the name of the Data Element in the definition (circular references).
State the definition in the singular .	Do not use emotive language , including persuasion.
Use language that defines what the Data Element is , and do not specify what it isn't.	Do not use numbers and special characters/symbols including underscores, commas, periods, and em-dashes.
Use casual business language , not technical jargon.	Do not use abbreviations . When needed, define a term once and include the abbreviation in parentheses immediately after. For example, "Mellon College of Science (MCS)" or "Tartan Scholars Fall (TSF)."
Be concise, precise, and specific .	Do not include procedural, usage, or rationale information in the definition.
Provide as much description as possible and include all relevant information .	

3. Other Information

Source of Definition

The Source of Definition is the internal organization or department that owns the Data Element being defined; it is **not** a form or a data system.

To provide a Source of Definition:

1. From the dropdown menu, **select the source where the definition of the data originated**.

Tip: Do not create a new source. If a source that you need is not available, contact the Data Steward for your definition's domain.

Classification

Classification indicates the **level of sensitivity of the data being described** by the definition.

Classification may determine who is assigned to steps in the workflow.

For example: If you are defining a term that you've classified as sensitive, the review of that term may be limited to people in specific user groups who are tasked with assessing sensitive data definitions.

Tip: Make sure your classification codes align with the [Guidelines for Data Classification](#)

To provide a Classification:

1. From the dropdown menu, **select the Classification code that best suits your definition.**

4. Technical Definition

Technical Definitions describe **how to obtain required data from a particular data system.**

Technical Definitions can be viewed in the [Data Cookbook](#) but are not visible through the [Data Glossary](#).

The Technical Definition is comprised of three parts. Follow the steps below to complete each part.

Tip: A report writer or administrator can help you obtain this information.

To create the Technical Definition:

1. From the **Data System** drop-down menu, select the system you identified in the [Definition Name](#) (if applicable).
2. *(Optional)* From the **Time Context** drop-down menu select the appropriate **time sensitivity**. For instance, a definition may be limited to a specific fiscal or financial aid year or an academic term.
3. Enter the technical details of the definition into the open text field, including:
 - Location of the Data Element in the Data System, written in the following format:
Table Name.Field Name
 - Data Element type
 - Data table
 - Data field

- Data type

If applicable, the open text field should also include the following information:

- **Data calculation or query information**

If the Data Element is the result of a calculation or query, include that query in the description and reference source tables and fields.

For example: to get the CostShare, you would have to look in two fields: CostShare (SPARCSProposals) + TotalCost (SPARCSProposals) = ProjectCost (SPARCSProposals)

- **Data System source name**

If the Data Element is a feed from another system of record, list the system in the description and the Data Element's name as it exists in that system.

For example: Employee Number in Oracle EBS originates as HRID in Workday.

- **External source information**

If the Data Element is from an **external source, describe the source in the description.**

- **Duplicate system information**

If a **Data Element is found in more than one system**, include multiple Technical Definitions for the definition.

For example: Student Address Type exists in the Student Data Warehouse (SDW) and the Student Services Suite (S3), so you must provide two Technical Definitions describing the Data Element in each system.

5. Tags

Tags are searchable keywords and phrases that can be shared across definitions. You can add a tag to a definition to indicate that it's associated with a certain project or has a certain characteristic (e.g., is FERPA protected).

To add a Tag to your definition:

1. Enter the appropriate tag(s) into the text field, separating multiple tags with commas.

For example: FERPA, COVID-19, Commencement.

6. Functional Area

A Functional Area is the topic or university entity such as Human Resources, University Advancement, the University Registrar, and others.

The Functional Area determines the definition's approval workflow and permissions for viewing and editing.

To specify a Functional Area for your definition:

1. From the **Functional Area** drop-down, select the area most relevant to your definition.