



Data Documentation Guide

Version 1.0

Document Control

| Version | Date | Description |
|---------|-----------|--|
| 1.0.0 | 8/12/2024 | Major revision. Updated the purpose and function of this document from a collection of information about the different data documents to a guide for completing the other DDP documents. |

Contents

| | |
|---|----|
| Introduction | 1 |
| Data Catalog - (DDP-2) | 2 |
| Corporate Data Element List - (DDP-3) | 5 |
| Data Element Standards - (DDP-4) | 8 |
| Data Diagrams - (DDP-5) | 10 |
| Dataset Metadata - (DDP-6) | 12 |
| Business Glossary Terms - (DDP-7) | 14 |
| Data Dictionary - (DDP-8) | 16 |
| Business Rules - (DDP-9) | 18 |
| Data Quality Management Plan - (DDP-10) | 20 |

Introduction

This document is a guide to creating and maintaining the data documentation standard of Caltrans. This is an integral part of Data Governance which is a standard process to improve data quality, documentation, interoperability, and access. This document describes and references several standard data documentation deliverables (DDP 2 through 10) that should be reviewed and completed, as necessary, for your type of data or system. See *Table 1* for a list of required and optional data documents.

- **Data Catalog** (DDP-2) – current and proposed corporate database and dataset entries for the Caltrans data catalog.
- **Corporate Data Element List** (DDP-3) – list of data elements proposed or approved for designation as “corporate data” that are designated authoritative sources of these elements.
- **Data Element Standards** (DDP-4) – existing or proposed data element standards that apply to data within the defined scope.
- **Data Flows** (DDP-5) – context and lineage diagrams describing how data flows across systems.
- **Dataset Metadata** (DDP-6) – documentation of dataset(s) packaged for end user consumption.
- **Business Glossary Terms** (DDP-7) – definitions of business terms represented by data entities or attributes within the defined scope.
- **Data Dictionary** (DDP-8) – detailed information about each data element (expands information provided in dataset metadata).
- **Business Rules** (DDP-9) – rules that provide extended documentation of individual data elements or validate data element or dataset contents.
- **Data Quality Management Plan** (DDP-10) – documentation of quality control and quality assurance measures applied to data as it is created/collected, maintained in systems, and exported for reporting or other purposes.

Table 1 - Required Standard Data Documentation

| Document Title | Doc. # | ● = Required ○ = Optional |
|------------------------------|--------|------------------------------|
| Data Documentation Guide | DDP-01 | n/a |
| Data Catalog | DDP-02 | ● |
| Corporate Data Element List | DDP-03 | ○ |
| Data Element Standards | DDP-04 | ○ |
| Data Flows | DDP-05 | ○ |
| Dataset Metadata | DDP-06 | ● |
| Business Glossary Terms | DDP-07 | ○ |
| Data Dictionary | DDP-08 | ● |
| Business Rules | DDP-09 | ○ |
| Data Quality Management Plan | DDP-10 | ● |

Data Catalog - (DDP-2)

The Caltrans Data Catalog is a listing of Caltrans' current and proposed corporate databases and datasets. It is used to meet several needs:

- To enable staff to search for and discover data assets.
- To facilitate compliance with the California Open Data Policy by providing a repository of datasets that can be considered for posting to the Open Data Portal.
- To enable Caltrans staff to identify current data resources and associated points of contact.
- To provide a resource for Caltrans Enterprise Data Stewards to understand existing data sources as they review proposals for acquisition of new data and systems.
- To serve as the authoritative source of Corporate Data Assets as approved by the Caltrans Enterprise Data Governance Board; and
- To track the Enterprise Data Stewards, Business Data Stewards and Data Custodians responsible for different Caltrans Data Assets.

The catalog will store information about Caltrans data assets that have been formally designated as corporate data, or that meet criteria for corporate data but have not yet been formally designated as such.

The catalog will store information about any dataset being considered for inclusion on an Open Data Portal.

The catalog will *not* store information about individual data elements, or the data itself, but will provide information on how to access the data. Data element information is documented in the [Data Dictionary](#).

The *Types of Corporate Data* are defined on the CTDATA [Corporate Data webpage](#)¹ and is replicated on the Corporate Data Fact Sheet. These resources describe the data types and definitions that must be chosen from when filling out the Corporate Data Category field in the template.

The *Subject Categories Taxonomy* table is from the Corporate Data Catalog which can be accessed on the [Onramp EDGG website](#)², as well. The User Guide is embedded on the main catalog page as well as the [Data Documentation Guide and Templates webpage](#)³ and contains the taxonomy list at the end of the document.

¹ <https://datagovernance.onramp.dot.ca.gov/corporate-data>

² <https://datagovernance.onramp.dot.ca.gov/>

³ <https://datagovernance.onramp.dot.ca.gov/data-documentation-package-and-templates>

Guidance and Process

1. Download the DDP-2.
2. Fill out the template completely.
3. Include links to any additional related background information.
4. Provide this information to the Caltrans Geospatial Data Officer (GDO).

References

[DDP-2](#) – Data Catalog Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp.

[CA Open Data Policy](#) – This includes information on the policy, as well as links to Technology Letter (TL) 19-01.

[Corporate Data Catalog](#) – Central repository for data assets at Caltrans.

[Subject Categories Taxonomy](#) – Table of data subject categories from which a corporate data asset can be assigned.

[Corporate Data Information](#) – This site has several informational resources, and it can be found on the EDGG website via Caltrans Onramp.

[Corporate Data Fact Sheet](#) – Simplified version of the Corporate Data information webpage.

Types of Catalog Entries

The following entries are distinguished in the Caltrans Data Catalog:

- **Enterprise Source Database** – an enterprise database (i.e., managed by Caltrans IT) containing source data. A source database is where data are actively maintained.
- **Enterprise Reporting Database⁴** – an enterprise database created specifically to provide a source for reporting, analysis or visualization. They store extracted, cleaned and transformed data sourced from other databases. Changes to data are only made through updates from the source databases. Examples include data warehouse tables, data marts, BI reporting universes.
- **File Data Source** – a spreadsheet or desktop-based database management system (DBMS) where data are maintained that is not managed as an enterprise data system by Caltrans IT.

⁴ Many Source Databases that are used to maintain data are also used for reporting. Only databases that are created specifically to integrate and transform data from other source systems should be categorized as Reporting Databases.

- **Dataset** – a collection of data elements prepared specifically for posting on a data portal or sharing with others. May be provided in various formats including comma-separated-values (CSV) file, spreadsheet, document, KML, shapefile, file geodatabase, and API.
- **Data Access Points** – online applications that provide interactive query capabilities, providing access to data not otherwise available through file downloads (for example, the Rebuilding CA map and tabular download webpage and DataLink for accounting data).

The catalog will store limited information for each entry. Additional metadata ([DDP-6](#)) and data dictionary ([DDP-8](#)) information is provided separately.

Corporate Data Element List - (DDP-3)

Corporate data elements, existing sources for those elements, and conformance with adopted presentation field name aliases are used for analytics and reporting. This is required, as a part of DGP-06, if sharing with others and/or if you intend to share data on the Enterprise Data Warehouse/and/or internally with other Caltrans business areas. The DDP-3 Template provides guidance for documenting data sources for corporate data.

In general, Corporate Data Elements should meet one or more of the criteria established for Corporate Data Assets:

- They represent master or reference data.
- They are required by control agencies.
- They are widely shared within or outside of Caltrans.
- They are procured or licensed data.

Corporate Data Elements and sources are documented and adopted as described in DGP-06 Designating SSOR and ARS for Corporate Data.

Effective governance of Corporate Data requires:

- Understanding where the data are maintained, and what sources should be used for reporting.
- Establishing clear and commonly understood data definitions.
- Establishing clear and easy to understand data element aliases for analysis and reporting.
- Establishing business rules for the data that provide the basis for assessing data validity.

Guidance and Process

1. Review DGP-06 – Designating SSOR and ARS of Corporate Data.
2. Review approved and posted Corporate Data Elements and ensure you will not be duplicating any entries.
3. Download the DDP-3 Template.
4. Form a Technical Working Group of appropriate subject matter experts.
5. Identify the source database(s) where the data element is entered and maintained.
6. Determine which of these source databases are considered the Source System of Record (SSOR). In general, there should be a single SSOR for each element. However, there will be exceptions. (For example, attributes for a project may have a different source system depending on what phase the project is in.)
7. Identify the authoritative reporting system containing the element, if applicable.
8. Review the field name and description for the data element as documented in the SSOR. Determine if the presentation name and abbreviation should be the

same as the data element field name and description in the source system of record and needs to be modified for understandability.

9. Complete the Corporate Data Element template to associate the source and reporting databases with the Corporate Data Elements.
10. Include links to any additional related background information in the comments field.
11. Work with Enterprise Data and Geospatial Governance Program staff to circulate the draft DDP-3 to technical stakeholders (as appropriate) and the Caltrans Enterprise Data Stewards (EDS) for review and comment.
12. Revise the draft DDP-3 as needed to respond to comments.
13. Work with Enterprise Data and Geospatial Governance Program staff to submit the final DDP-3 to the Caltrans Enterprise Data Governance Board for review and approval.
14. The Caltrans Enterprise Data Governance Board may approve the submittal or return it with comments to be addressed.
15. Repeat steps 5-14 to modify the proposal as needed.
16. Help put the standard into practice by communicating with colleagues and advocating for its use. The EDS will ensure that the business data stewards are informed, and the standard is posted online.

References

[DDP-3](#) – Corporate Data Element List Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp.

[DGP-06](#) – Designation Source Systems of Record and Authoritative Reporting Systems for Corporate Data posted on the EDGG website via the Caltrans Onramp.

[Corporate Data Catalog](#) – This is where Caltrans corporate data assets are identified.

Source Type Definitions:

- **A Source System (SS)** – a database in which the data element is entered or updated but is not the recommended place to obtain the data. For example, a non-authoritative system may be used to support an isolated business process and have no connection to other databases. Data in this system is not necessarily current.
- **A Source System of Record (SSOR)** for the data element. This is a source system where data are entered and updated that you would point to as the most authoritative place to obtain the data element. In some cases, there will be multiple SSORs for a data element – for example, project information may be sourced from one system during planning and another system during scoping and design.
- **An Authoritative Reporting System (ARS)** - this is a data warehouse table or other reporting database (e.g., Tableau data source) that has been designated as an authoritative source for reporting. Data are not modified within this system – data are pulled in from the SSORs and may be transformed to provide consistency

and ease of reporting. In cases where there are multiple SSORs for a data element (e.g., SSOR for a given data element changes over the project lifecycle), the Authoritative Reporting System may be the most reliable place to obtain the element.

Data Element Standards - (DDP-4)

Data governance initiatives can propose new Caltrans data element standards to address inconsistencies in data elements across multiple systems. Lack of consistency across data sources can create problems when there is a need to combine data from multiple systems for reporting. Lack of consistency also presents a barrier to integrating systems to improve efficiencies.

Additionally, listing out what existing data element standards your system complies with or will comply with in a future state is valuable. If your system does not comply with an existing data element standard, a plan should be developed to implement an enhancement to make it compliant. You may also want to consult with Enterprise Data and Geospatial Governance staff (add shared e-mail) for additional guidance.

Guidance and Process

Data element standards for individual data elements should include the following:

1. Definition of the data element – sufficient to avoid varying interpretations of what it means.
2. Value Domain/Reference Tables – including code list or source for the list of possible values if applicable. Please confirm with Enterprise Data and Geospatial Governance Program staff that you are using approved reference tables. Where a needed reference table doesn't exist, you may consider proposing one be developed and adopted for Caltrans.
3. Data type (e.g., numeric, character, date)
4. Scale/units of measure (e.g., Celsius or Fahrenheit for a temperature)

In addition, the following optional items can be considered, as appropriate:

1. Naming convention (for the physical database field name)
2. Positional accuracy (e.g., +/-40 feet)
3. Coordinate system
4. Reference(s) that provide further detail on how the data element is to be captured or calculated.

To create a new data element standard:

1. Create a problem statement describing the need for the standard.
2. Research existing data element standards to determine if there is an existing related standard that can be modified.
3. Form a working group involving individuals who have technical knowledge of how the data element is used within the different systems where it is stored.
4. Collaborate to draft a proposed standard that will address the identified need and can be feasibly implemented.
5. Work with Enterprise Data and Geospatial Governance Program staff to circulate the draft standard to technical stakeholders (as appropriate) and the Caltrans Enterprise Data Stewards (EDS) for review and comment.

6. Revise the draft standard as needed to respond to comments.
7. Work with Enterprise Data and Geospatial Governance Program staff to submit the draft standard to the Caltrans Enterprise Data Governance Board for review and approval.
8. The Caltrans Enterprise Data Governance Board may approve the standard or return it with comments to be addressed.
9. Repeat steps 4-8 to modify the draft standard as needed.
10. Help put the standard into practice by communicating with colleagues and advocating for its use. The EDS will ensure that the business data stewards are informed, and the standard is posted online.

References

[DDP-4](#) – Data Element Standards Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

A list of proposed and accepted Caltrans data element standards can be found here: [Data Elements | Enterprise Data and Geospatial Governance \(ca.gov\)](#)

[Data Element Standards Form](#) – This online form allows data stewards and users to nominate and help develop Caltrans data element standards to address inconsistencies in data elements across multiple systems.

Data Diagrams - (DDP-5)

Several types of data diagrams can be produced to document how data are entered or loaded into different information systems, and how they may be transformed along the way. These diagrams are valuable supplements to tabular metadata – they can be used to support data quality improvement efforts and data change control processes. They also help data stewards, custodians and users gain a common understanding of how data are produced and managed.

It is recommended that each data governance initiative produce (at a minimum) a high-level data flow diagram – also called a context diagram.

Guidance and Scope

Data flow and lineage diagrams are contextual diagrams describing how data flows across systems. There is no standard documentation for data flow diagrams at this time, but some examples are provided for reference in the link for the DDP-5 samples below. These diagrams can cover multiple data assets.

Process

1. Download the DDP-5 Sample Context Diagrams to get a good idea of how to create one.
2. Determine a process to create and validate the diagrams, including who will be involved and what sources you will use.
3. Create the Data Diagram/s.
4. Provide this information to the Caltrans Geospatial Data Officer (GDO) and keep a copy with all other system documentation.

References

[DDP-5](#) – Data Flows examples found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

Data Flow Diagrams

Data flow diagrams are a well-established method for documenting how data moves from people/roles into and out of processes and data stores. A process is something that takes data as input and produces data as an output. Examples of processes are permitting, order processing, and data validation. A standard data flow diagram shows:

- External Entities - the people/roles/other systems that generate and/or receive information.
- Data Stores – the databases or files that store information.
- Processes - the processes that occur in the system to manipulate the information.
- Data Flows - the information that enters and leaves the system and flows across processes and data stores within the system.

Data flow diagrams can be prepared at different levels of detail. The highest-level diagram is called a "context diagram". A context diagram is useful for showing what information is flowing in and out of a system, and the external sources and recipients of the information. More detailed data flow diagrams show information flows within a system's constituent processes and data stores.

A tutorial on preparing data flow diagrams is available [here](#).

A sample data flow diagram is available [here](#).

Data Use Diagrams

A variant of a data flow diagram is a Data Use Diagram. A data use diagram shows how different business units use data from a selected information system.

A sample Caltrans data use diagram is available [here](#).

Data Value Maps

A data value map identifies types of users and uses for the data, and traces these to data products and data sources. An example data value map is available [here](#).

Lineage Diagrams

Lineage diagrams trace the path of one or more data elements from their original point of entry to their final destination(s). They may include documentation of transformations performed along this path. Lineage diagrams are helpful for identifying and diagnosing data quality issues that may cascade across multiple databases as data are transferred from one system to another (via manual re-entry or automated processes). They are also helpful for analyzing the potential impacts of a change to a data element's structure or definition within one data store on downstream data stores or reports. One common way to approach a lineage diagram is to begin with a critical report and then trace the elements on that report backwards to their origin. Because lineage diagrams can be complex, they are typically created and managed with specialized tools.

Dataset Metadata - (DDP-6)

A dataset is defined as a collection of data elements prepared specifically for posting on a data portal or sharing with others. Datasets may be provided in various formats including spreadsheet, document, KML, shapefile, file geodatabase, API.⁵

Dataset metadata provides information about a dataset as a whole. It is supplemented by Data Dictionary information for each data element in the dataset. Dataset metadata helps people to search for data of interest because the metadata elements can be used as filter or search criteria in a data catalog or data portal. Metadata also helps people to understand the purpose, derivation, and limitations of a dataset. It provides important information for determining whether a dataset is suitable to meet a particular information need.

Note that this guidance was developed for documenting collections of data elements assembled for distribution or reporting. It is not intended to be used for documenting every physical database table within a system.

Guidance and Scope

Dataset-level metadata is information about the dataset as a whole. For GIS datasets, some required metadata is also in the Data Dictionary ([DDP-8](#)). Make sure to review and provide all the required elements. Metadata should be created for all final datasets.

Process

1. Download the DDP-6 template.
2. Fill out the template.
3. Submit the completed DDP-6 template to the Caltrans Geospatial Data Officer (GDO) for review.
4. Post with the dataset if shared online or provide with a dataset as part of a data deliverable.

⁵ Note that the California Open Data Handbook uses a different definition of dataset: “any organized collection of data. The most basic dataset is composed of data elements in a table. Each column represents a particular variable. Each row corresponds to a given value of that column’s variable. 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. Dataset is a flexible term and may refer to an entire database, a spreadsheet or other data file, or a related collection of data resources.” The Caltrans definition is not inconsistent with this but is intended to make a clear distinction between datasets designed for sharing or distribution and databases (and their constituent tables) are part of transactional systems and not suitable for sharing in their current form.

References

[DDP-6](#) – Dataset Metadata Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

[DGP-02](#) – Caltrans Metadata Standard found on the Data Practices Document page on the EDGG website via the Caltrans Onramp.

Metadata Items

Caltrans has established the following recommended metadata elements to be provided for all datasets to be designated as Corporate Data Assets. These elements were selected for consistency with the guidelines in the [California Open Data Handbook](#) and the Federal Geographic Data Committee's (FGDC) recommended metadata standard.

Data classification terminology (e.g., Personal Identity Information (PII), confidential information, etc.) is defined per California Department of Technology [Data Classification Standard](#).

GIS Data

Descriptions of how to enter metadata are provided in the DDP-6. In ArcGIS Pro, the metadata required style is ArcGIS Pro Metadata Style ISO 19139 Metadata Implementation Specification GML 3.2. To define the metadata style in ArcGIS Pro, go to the Project settings and click on Options to find Metadata. You can set it as your metadata style there.

Make sure to start editing Metadata by going to the Catalog Pane and right clicking the feature class in the file geodatabase. This will give you access to the metadata in the data source which enables sharing. If you access the metadata by right clicking the layer in the Contents pane of a map, it will not save in the data source and be stored when shared, pasted, or dragged to a new location.

There are three categories – Overview, Metadata, and Resource - with subcategories organizing the information within. When you click on a subcategory the data entry forms are visible.

Make sure to review the Data Dictionary (DDP-8) as there are GIS metadata requirements listed there for the Fields.

Business Glossary Terms - (DDP-7)

A business glossary provides a central location providing vetted, shared definitions of common business terms. While a data dictionary includes descriptions of the meaning of each data element, a business glossary defines the core concepts that data may represent. For example, a business glossary might include terms like "State Highway System", "State Highway Operations Program", "STIP Project", "culvert" and "gore point".

Because there are many independent glossaries of terms maintained by Caltrans business units, building a unified Caltrans business glossary will be an incremental process. Initially, the glossary will contain those terms that are important for understanding concepts underlying designated corporate data assets.

Guidance and Scope

Business glossary terms represent the key business concepts within the scope of the dataset it represents. Business glossary terms should be created using the standard Caltrans business glossary template or (when available) added directly to the Caltrans Business Glossary.

Process

1. Download DDP-7 template.
2. Complete the template to compile business glossary terms.
3. Provide this information to the Caltrans Geospatial Data Officer (GDO) for review.
4. Create entry in Caltrans Business Glossary.
5. Update status of term to Approved when adopted by CTDATA.

References

[DDP-7](#) – Business Data Glossary Terms Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

Business glossary terms are listed in Caltrans [Business Data Glossary](#)

Style Guidelines for Writing Definitions

1. Don't include the term you are defining within the definition.
2. Name the general class of objects to which the concept underlying the term belongs and identify distinguishing characteristics of the concept. (e.g., "A state highway is a roadway owned and maintained by Caltrans.")
3. Be succinct – don't include extraneous information not essential to understanding the term.
4. Avoid unnecessary jargon – make the definition understandable to a general audience.

5. Use active voice and first person singular unless the term being defined is plural or a collective noun.

Do rely on authoritative sources for definitions (e.g., legislation or existing glossaries) but make sure the source is in the public domain before copying the definition verbatim. Paraphrase definitions that may be copyrighted.

Data Dictionary - (DDP-8)

A data dictionary provides information about the different data elements or fields in a dataset. It serves several different purposes:

- Capture and preserve information about data elements to maintain and sustain a common understanding of the data.
- Help future data users to understand the meaning of each data element.
- Articulate business requirements to be used by database developers to design new databases or modify existing databases.
- Communicate current database design characteristics to application and report developers.

If a physical database already exists, some the data dictionary items can be extracted from the database schema and used as a starting point for documenting the complete set of items listed below. This guidance can be used to produce a data dictionary for an entire database by table as well as documenting the data elements within a dataset – which is defined as a collection of data elements prepared specifically for posting on a data portal or sharing with others.

Guidance and Scope

Data dictionaries should be created using the standard Caltrans data dictionary template. Elements required in GIS datasets are also present here. All data assets need a data dictionary to provide information about the different data elements or fields.

Process

1. Download the DDP-8 template.
2. Fill out all required and/or optional fields.
3. Send completed DDP-8 to the Caltrans Geospatial Data Officer (GDO) for review.
4. Post with the dataset if shared online or provide with a dataset as part of a data deliverable.

References

[DDP-8](#) – Data Dictionary Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

[DGP-03](#) – Caltrans Data Dictionary Standard found on the Data Practices Documents page on the EDGG website via the Caltrans Onramp

Data Dictionary Items

The data dictionary items in DDP-8 are to be included for documenting corporate data. Mappings to Caltrans ISO 19139 GML 3.2 geospatial metadata elements (GIS datasets) and the California Open Data elements are provided in the DDP-8 Template. Some elements are a crucial part of a GIS datasets metadata and should always be

included. For data in a GIS format, items marked with an asterisk in the template are managed from within the GIS software and should not be edited using a metadata editor.

The worksheet “List of Values” must be completed if the data has defined values (whole words or codes) or uses a reference dataset as a source (can be a link to a URL). Some List of Values examples are city names, county abbreviations, culvert sizes, pavement types, employee classifications, etc.

Business Rules - (DDP-9)

Business rules provide the foundation for data quality management. They can be used to build in data validation logic within data entry applications, and to develop procedures for data validation and cleansing once data are already in hand. They also are used to provide guidance for manual data review.

Data entry and validation logic for individual data elements should be based on information provided in the data dictionary. This information includes allowable and expected ranges (min and max) for numeric and date data elements and lists of values for enumerated and coded data elements. Usage notes for data elements should also be included in the data dictionary to supplement the more general-purpose field descriptions and provide extended technical documentation of complex logic for calculated, composite or overloaded fields.

Business rules that go beyond validation of individual data element values are described below.

Guidance and Scope

If your dataset has a business rule that would aid in data validation or provide guidance for data review, this should be completed. Some business rule types are Corporate, Geo-Processing and Linear-Referencing, Cardinality, Dataset Level, and Usage Notes/Descriptions.

Process

1. Download the DDP-9 template.
2. Fill out required and/or optional fields.
3. Send the completed DDP-9 template to the Caltrans Geospatial Data Officer (GDO) for review.

References

[DDP-9](#) – Business Rules Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

Types of Business Rules

The following types of business rules are to be considered and included in the business rules catalog:

| <u>Rule Type</u> | <u>Rule Description</u> |
|--|--|
| Conditional Rules | Conditional rules applied to one data element based on values of other data elements. |
| Geo-Processing and Linear-Referencing Rules | Rules applied to a data element or the record itself that utilize the location of a record to validate data entry. |

| | |
|---------------------------------|--|
| | These rules may validate data entry or record creation based on a GIS feature or information embedded within other common location referencing information. |
| Cardinality Rules | Rules applied to and describing the relationships between the data entities within a given dataset (one to one, one to many, etc.). These rules are used to inform or validate the relational model. |
| Dataset Level Rules | Rules applied to the dataset at large, typically applied to aggregation of values across records for comparison against other attributes of the dataset or an expected value. |
| Usage Notes/Descriptions | Details of calculations or other business logic that is too detailed to include in the data dictionary but important for understanding and validating the data element. |

Data Quality Management Plan - (DDP-10)

Guidance and Scope

See the Caltrans Data Quality Management Plan (DQMP) Template for guidance on creating a plan that describes data quality objectives, the current level of data quality (if known), current and future planned quality management practices, and quality tracking and communication strategies. The DQMP should be a living document, updated regularly. Data quality management plans should cover data collection, data being maintained in a database as part of a primary application or system, and for extraction for reporting purposes.

Process

1. Download the DDP-10 template.
2. Fill out all required fields.
3. Send the completed DDP-10 template to the Caltrans Geospatial Data Officer (GDO) for review.

References

[DDP-10](#) – Data Quality Management Plan Template found on the Data Documentation and Templates page on the EDGG website via the Caltrans Onramp

[DGP-01](#) – Data Quality Management Plan Implementation found on the Data Practices Documents page on the EDGG website via the Caltrans Onramp

More Resources

CTDATA hopes that this document clarifies the process of completing data documentation in data governance implementation. Additional resources are also available on our [intranet site](#).⁶ If you have questions about any of the content within this document, or want to learn more about data governance, please email us at CTDATA@dot.ca.gov

- Caleb McCallister, PLS
Enterprise Data and Geospatial Governance Program Manager



- End of Document -

⁶ <https://datagovernance.onramp.dot.ca.gov/>