California Department of Transportation **Enterprise Data and Geospatial Governance Practice**	**NUMBER:** DGP-01	**DATE ISSUED:** AUGUST 11, 2021
SUBJECT: Data Quality Management Plan (DQMP) Implementation	**REFERENCES:** Caltrans Data Documentation Package Enterprise Data and Geospatial Governance (EDGG) Fact Sheets	

BACKGROUND AND PURPOSE

Caltrans has adopted a set of core data principles[1], based on those established by the American Association of State Highway and Transportation Officials (AASHTO). These principles recognize that data is an asset that needs to be maintained – much like physical assets. Good stewardship of data involves ensuring that data is of sufficient quality to serve its intended purposes.

Data is used in all areas of the transportation decision-making process during the project lifecycle (planning, design, construction, maintenance and operations). Furthermore, data is increasingly being used externally by citizens and customers to facilitate their personal decisions, economic personalities and personal development, and by stakeholders to assess the aggregate performance of a transportation organization. Significant human and system resources are consumed in the collection, analysis and dissemination of data irrespective of the quality. It is essential that the most effective use of public funds be achieved through appropriately directed attention to data quality and the procedures to realize and maintain that quality.

DATA QUALITY:

"Quality data" means data is fit for their intended purposes. Several aspects of data quality include accuracy, timeliness, and completeness. Data may be highly accurate but produced too late to be of value to users. Conversely, data may be provided in real time, but lack credibility. Usability is also an important consideration, and the ease with which an end user can access and utilize the data to meet their need. Data may be accurate and timely, but if it is not accessible or requires specialized expertise to transform into a useable form, then it does not meet intended uses.

PRACTICE:

For Caltrans to attain and maintain data quality, each business area shall complete and implement a Data Quality Management Plan (DQMP) for each of their corporate datasets, database systems and corporate data over its life cycle as follows:

1. Each business area shall complete each section of the DQMP template, including data quality objectives and targets, potential sources of error, actions to improve metrics, compliance with all laws, regulations, and policies and update frequency. Some datasets may require preparation by staff with professional credentials such as bridge or surveying data.
2. For business areas with existing DQMPs or data quality documentation, the DQMP template shall be used to focus on other data quality areas not currently covered by existing practices or to document other areas for data quality improvement.
3. For business areas with existing data quality practices, those practices shall be documented in the DQMP template in addition to methods to improve data quality objectives.
4. Business areas are highly encouraged to use data quality tools or automated methods to verify data quality. For example, database constraints shall be used to prevent invalid data from being entered into the system.
5. Manual checking of data is discouraged since it is time-consuming, prone to human error and not scalable. Where automated checks are not feasible, manual checks will be considered sufficient.

[1] AASHTO Data Principles: https://datagovernance.onramp.dot.ca.gov/data-principles-and-benefits

APPLICABILITY:

This practice document shall apply to:

1. Source System of Record (SSOR)
2. Authoritative Reporting System (ARS)
3. Corporate Datasets
4. Corporate Data Lifecycle
    a. Field Data Collection
    b. Data Procurement (Vendor-Provided)
    c. Data Entry and Maintenance
    d. Data Extraction and Loading

The DQMP shall apply to existing data and new data upon entry into data repositories. Validation checks such as input field validation, input maps and look up values shall be implemented whenever possible. DQMP serves as a living document which should be implemented with existing information and improved over time.

For corporate datasets, a DQMP shall be implemented and maintained by the business area responsible for the corporate dataset and be reviewed annually. Proposed improvement identified during the annual assessment, or more frequently as needed to ensure that the data meet the intended purpose, will be compiled into an Annual DQMP Improvement Plan and submitted to the responsible Division Chief for approval.

For data collection, a DQMP shall be completed as part of the scope of work and will cover how the data is collected in the field such that it meets the data quality standards necessary for its intended use.

TIME FRAME:

This practice document shall take effect immediately for corporate datasets. Where immediate compliance is not possible, the DQMP shall be updated during the next scheduled data migration or upgrade cycle.

DQMPs shall be completed for existing systems and new SSOR and ARS within 1 year upon nomination by the Enterprise Data and Geospatial Governance Board and Enterprise Data Stewards Committee. For new corporate datasets, DQMPs shall be completed upon nomination.

DQMP development and implementation shall be coordinated with other efforts, upgrades and migrations within each business area. Exemptions to this time frame shall be requested in writing from the Enterprise Data and Geospatial Governance Program.

ROLES / RESPONSIBILITIES:

Enterprise Data Stewards, Business Data Stewards and Data Custodians are responsible for developing and maintaining DQMPs for the corporate database system, corporate datasets and field data collection within their business area. Staff within each business area are responsible for implementing DQMPs and improvement plans.

The Division Chief for each business area shall approve the DQMP and Annual DQMP Improvement Plan prior to their adoption. The Division Chief shall support DQMP implementation by providing sufficient resources, guidance to staff and encourage collaboration with other business areas.

IMPLEMENTATION:

DQMP implementation shall start with corporate data and then expand to data collection and individual flat file systems. Implementation should include an as-is assessment of existing data quality.

Each business area shall establish metrics in their DQMPs which can be reasonably achieved. Objectives shall be proposed and met over time, then adjusted for improvement. Each business area shall work to improve their metrics over time.

Data quality objectives and metrics may vary between systems. Hence, each business area shall consider their existing data quality processes to develop objectives and metrics. For large, complex database systems, staff are encouraged to start their DQMP implementations early. For business areas that store data in flat files, staff are encouraged to complete and implement DQMPs.

Each business area is expected to complete their DQMPs and work with appropriate staff on its implementation. An action plan may be needed to organize implementation by task to achieve DQMP objectives and metrics within estimated complete dates. The business area shall prioritize DQMP tasks and determine which ones are mandatory or optional.

PERFORMANCE TARGETS:

DQMP completion shall be tracked and reported to the Enterprise Data Stewards Committee and approved by the Enterprise Data Governance Board annually.

ATTACHMENTS:

1. DQMP Template
2. DQMP Example
3. EDGG Roles and Responsibilities Fact Sheet
4. EDGG Glossary Fact Sheet
5. EDGG Corporate Data Fact Sheet
6. DQMP Metrics and Objectives Worksheet

APPROVAL:

Chad Baker Date: August 11, 2021

Chair, Enterprise Data Governance Board or Designee
California Department of Transportation