

California Department of Transportation <b>Enterprise Data and Geospatial Governance Practice</b>	<b>NUMBER:</b> DGP-02	<b>DATE ISSUED:</b> AUGUST 11, 2021
<b>SUBJECT:</b> Metadata (dataset) Standard	<b>REFERENCES:</b> Caltrans Data Documentation Package DDP-6 Caltrans Dataset Metadata Template CA Open Data Handbook DES-04-R1 Location – Coordinate Specification	
<b>BACKGROUND AND PURPOSE:</b> <p>Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource (National Information Standards Organization (NISO) 2017, ISBN: 978-1-937522-72-8). The standardization of a metadata schema and definitions allows for uniform access to fundamental information about an information or data resource. This also creates a structured environment allowing for consistent and reproducible searching, publishing and harvesting in a manner irrespective of the platform being used. Metadata enables greater discovery and identification of datasets that meet a given business need and ultimately enables users of an information resource to determine if the data within is fit-for-use in their analysis.</p> <p>The purpose of this practice document is to identify the dataset metadata standard for Caltrans for corporate datasets that must be maintained and included when data is shared with data consumers.</p>		
<b>DEFINITIONS:</b> <p>Data consumer – Any recipient of a dataset.</p> <p>Corporate data – Data assets that have been prioritized for data governance based on their value to Caltrans and its customers and stakeholders, and where lack of attention to data standardization, documentation, quality management and coordination introduces inefficiencies and risks. Please visit the following webpage for full details:  <a href="https://datagovernance.onramp.dot.ca.gov/corporate-data">https://datagovernance.onramp.dot.ca.gov/corporate-data</a></p>		
<b>RESOURCES:</b> <p>Caltrans Data Documentation webpage. This webpage includes the Data Documentation Package and DDP-6 Caltrans Dataset Metadata Template documents.  <a href="https://datagovernance.onramp.dot.ca.gov/data-documentation">https://datagovernance.onramp.dot.ca.gov/data-documentation</a></p> <p>CA Open Data Handbook. The online document explains the requirements around publishing open data and includes documentation on the state open data metadata standard.  <a href="https://handbook.data.ca.gov/">https://handbook.data.ca.gov/</a></p> <p>Understanding Metadata: What is Metadata, and What is it For?: A Primer  <a href="http://www.niso.org/publications/understanding-metadata-2017">http://www.niso.org/publications/understanding-metadata-2017</a></p>		

**PRACTICE:**

Dataset metadata shall be completed and maintained for all Caltrans corporate datasets. The metadata elements that must be completed are described in the Data Documentation Package Appendix E and DDP-6 Caltrans Dataset Metadata Template. Individual metadata elements are identified as required, if applicable or optional.

*Geographic Information System (GIS) format datasets* (shapefile, file geodatabase, web feature service, etc.) GIS format datasets shall use the ISO 19139 GML 3.2 metadata specification for storing metadata. Dataset metadata shall be directly managed as part of the dataset itself using a metadata editor such as Esri's ArcCatalog. Where it is impossible to manage the metadata in this fashion, the metadata must be provided as an ISO 19139 GML 3.2 compliant XML file or as a CSV export from the DDP-6 template.

*Non-GIS format datasets* (CSV, XLS, XML, JSON, etc.)

Non-GIS format datasets shall use DDP-6 for recording dataset metadata and exported as a CSV file.

All corporate dataset metadata shall be maintained over the entirety of its lifecycle and shall be delivered with all corporate datasets to data consumers.

**THIRD-PARTY SOURCED DATA:**

For externally sourced data where the responsible entity is unable or unwilling to create compliant metadata but for which Caltrans reuses, reshapes, and/or enhances the dataset, the responsible Business Data Steward must ensure that compliant metadata is created.

**TIME FRAME:**

This practice document shall take effect immediately and existing published corporate datasets shall have compliant metadata created within 30 days. If additional time is required for compliance of existing published corporate datasets, business areas may request a time extension via email to the Geospatial Data Officer.

Newly created corporate datasets shall have metadata created as part of the overall dataset creation process.

**IMPLEMENTATION:**

Caltrans Business Data Stewards and Data Custodians should lead efforts to create metadata for all existing corporate datasets and ensure all new datasets created have metadata created concurrently. When datasets are shared to the open data portal, emailed, or distributed through webpages or any other means, the associated dataset metadata shall be provided.

**ROLES AND RESPONSIBILITIES:**

Enterprise Data Stewards, District Enterprise Data Governance Liaisons, Business Data Stewards, Data Custodians and other identified staff shall lead efforts to modify or augment existing practices as necessary to implement the requirements of this Practice document.

Division Chiefs and Deputy District Directors shall, where applicable, direct and support staff to create, maintain and distribute metadata for their corporate datasets.

Consultants working on projects and/or are under contract with Caltrans shall follow the requirements of this Practice document.

When entering into agreements or contracts with an external entity, the requirements of this Practice document should be addressed as part of the negotiation process.

**APPROVAL:**

*Chad Baker*

Date: August 11, 2021

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