

Publishing Metadata Guide for DATA.CA.GOV

Metadata Field	Required/Optional	Field Description
Title	Required	<ul style="list-style-type: none"> Choose a name for your dataset that tells users what the topic of the dataset is. Do not use CA, California or dates in the Title.
Description	Required	<ul style="list-style-type: none"> This is a plain English description that will display below the name of the data table. Write a summary paragraph telling us what the data table contains. The first few sentences are key. Include related legislation info if applicable. Include acronyms that people might look for here, but avoid acronyms in your first few sentences.
Tags	Required	<ul style="list-style-type: none"> Descriptive keywords or phrases that users will search for to find your data resources. Separate each keyword by commas. Try to include at least five descriptive tags. Both general and specific terms are useful. Programs and acronyms, e.g. California State Library (CSL), can make finding your data resource easier.
Groups	Required	<ul style="list-style-type: none"> Also known as the publisher The agency, group, department, board, or commission that publishes the data resource. Choose from the drop down list.
Topics	Required	<ul style="list-style-type: none"> Also known as Category. Choose the most appropriate topic from the current list at https://data.ca.gov.
License	If-Applicable	<ul style="list-style-type: none"> Most often Public Domain. Any restrictions on copying, sharing, using, etc. your data must be disclosed.
Author	Optional	<ul style="list-style-type: none"> The agency, group, department, board, or commission that authors the data resource and has ultimate responsibility for the creation of the data. Will often but not always be the publisher of the data.
Spatial/ Geographic Coverage	If-Applicable	<ul style="list-style-type: none"> The geographical area the data table covers (e.g. statewide versus a sub-state region like the Bay Area). Specification should include a named area and may include geographic coordinates.
Frequency	Required	<ul style="list-style-type: none"> How often do you intend to publish or update the data resource on Data.ca.gov? E.g. Annually, quarterly, monthly.
Temporal Coverage	If-Applicable	<ul style="list-style-type: none"> Start date and End date for the data in your data resource.
Granularity	Optional	<ul style="list-style-type: none"> Tell us the most specific that data the in your resource gets. Often measured in geography (county, census tract) or time (monthly or daily data).
Data Dictionary Type	Required	<ul style="list-style-type: none"> If the link below is not an HTML file, the file type for the data dictionary (most often a PDF).
Data Dictionary	Required	<ul style="list-style-type: none"> HTML link to the data dictionary itself.
Program Contact Name	Required	<ul style="list-style-type: none"> The specific group inside the agency, department, board or commission that produces the data that can best answer questions about the data.
Program Contact Email	Required	<ul style="list-style-type: none"> Enter a generic e-mail address for the program referenced above. (e.g. answers@library.ca.gov)
Public Access Level	Required	<ul style="list-style-type: none"> Whether this info could ever be made public. (Public, Restricted, Non-Public)
Rights	Required	<ul style="list-style-type: none"> If you entered anything but Public above, you must explain any use restriction on the data. Additionally, you can also include usage/research/collaboration instructions: Short text or link to a document that describes how the data can be used, research Ideas and/or possible collaborations based on this information that may interest external researchers. If not applicable, please enter "No restrictions on public use".
Homepage URL	If-Applicable	<ul style="list-style-type: none"> URL for the page on your website that has useful information about the data resource or the group that updates it. It's a webpage where the user can download this and additional pertinent information.
Data Standard	If-Applicable	<ul style="list-style-type: none"> This is used to identify a standardized specification the dataset conforms to. It's recommended that this be a URI that serves as a unique identifier for the standard. The URI may or may not also be a URL that provides documentation of the specification. A technical description of the data. E.g. CSV, XML standards, SHP, or JSON.
Language	If-Applicable	<ul style="list-style-type: none"> Most often English.
Additional Information	Optional	<ul style="list-style-type: none"> Additional information is a field that allows you to enter free form metadata in key-value pairs: a key, which is a unique identifier for some item of data, and the value, which is the data that is identified. You must include at least two keys-value pairs: one with the "Limitations" key and one with the "Data_Methodology" key. Limitations/Exclusions: Must include the following required text: "Use of this data is subject to the CA.gov Conditions of Use and any copyright and proprietary notices incorporated in or accompanying the individual files." <ul style="list-style-type: none"> This may be followed by a brief description of any limitations on these data or of exclusions to their use not otherwise covered above in "Rights". Data Methodology: Short text or link to a document with explanation of the data collection methodology, which may include survey tools, post-collection methods for control and cleaning and notes on sampling response and errors. This is where you should document any known issues in using the data. May also be uploaded as an additional PDF.
Related Content	Optional	<ul style="list-style-type: none"> Enter secondary source(s) info: If your data resource is partially made from other data sources, please give descriptive name(s), and/or URLs, of resource(s) from which the data table is derived.