



OPEN DATA POLICY AND STANDARDS

December 2023

Version 2.0

Table of Contents

EXECUTIVE SUMMARY	3
1 PURPOSE	4
2 ROLES AND RESPONSIBILITIES	4
2.1 Department of Information Technology	4
2.2 Data Custodians	4
2.3 Data Owners	4
2.4 Open Data Council	5
3 GENERAL DATA SPECIFICATIONS	5
4 METADATA	5
4.1 Metadata Requirements	5
5 DATA MAINTENANCE	7
6 QUALITY ASSURANCE / QUALITY CONTROL (QA/QC)	7
6.1 Data Owners/Custodians	7
7 DATA SECURITY	
8 PLATFORM SECURITY	8
9 ORGANIZATION	8
9.1 Organization of Data	8
10 PROCEDURES FOR PUBLICATION OF DATA	10

EXECUTIVE SUMMARY

Data serves as one of the State's most valuable assets and has become a vital component for many organizations. It is also a critical piece of the decision-making process. Maryland enacted the Open Data Act in 2015 to make state data open and easy to find, access, and use by the public, through open data portals.

Maryland's open data portals contain nearly 2,000 datasets of agency data. The intent of the Portals is to offer all of Maryland's organizations, both public and private, a resource to help them maintain transparency, hold themselves to account, and make optimal decisions based on the best possible information. County and association partners contribute additional datasets so that Maryland's open data portals offer centralized access to public data in machine readable format. The format allows Portal customers to explore, export, feed, filter, and visualize all of the datasets on the Portal. With easy access to and flexible manipulation of datasets, the Portals offer users access to clean, standardized, and relevant data that is critical to them.

The specifications and guidelines in this Data Management Plan will improve data consistency and availability of information. It will ensure that all levels of government and the public have access to the most up-to-date information, reduce or eliminate overlapping data requests and redundant data maintenance, and ensure metadata is consistently created.

This Data Management Plan contains the following sections:

- *Roles and Responsibilities* - Entities involved with the Open Data Portal and their responsibilities.
- *General Data Specifications* - Data specifications that pertain to all open data. The goal is to adhere to the recognized national and international data standards.
- *Metadata* - Datasets included on the Open Data Portal are the basis for many applications and services. This section contains instructions for Open Data Portal dataset metadata requirements.
- *Data Maintenance* - The data owner/custodian will be responsible for ensuring the data on the Open Data Portal is kept up-to-date.
- *Quality Assurance/Quality Control* - Data owners/custodians will be responsible for QA/QC of these data. Recommended minimum steps are provided.
- *Data Security* - Security will conform to MD Department of Information Technology's Information Security Policy.
- *Organization* – Open Data Portal content will be grouped in categories.
- *Procedures for Publication of Data* - Instructions on preparing and submitting data for inclusion on the Open Data Portal.

1 PURPOSE

The policy and standards within this document will improve data consistency and availability of information.

The Open Data Portal is a central store of data, giving the public and staff from all organizations, both public and private, access to the most recent data that can be critical to their success.

The Open Data Portal central data store reduces costs and improve the effectiveness of agency efforts by:

- Ensuring that all levels of government and the public have access to the up-to-date information from other agencies;
- Reducing or eliminating:
 - Overlapping data requests; for example, multiple agencies contacting SHA for the latest road centerlines;
 - Redundant data maintenance
- Ensuring metadata is consistently created, maintained, understood and used
- Data services can be displayed by the consumer agency using the application and format of its choice.

2 ROLES AND RESPONSIBILITIES

This section contains a list of entities involved in management of the Open Data Portal and their responsibilities.

2.1 Department of Information Technology

- DoIT is responsible for management of the Open Data Portal.
- DoIT will be the default custodian on any data layers that are ambiguous, data layers that do not have a clear custodian.

2.2 Data Custodians

- The data custodian is not necessarily the data owner. For example, MDA may repackage USDA data but they are not considered the data owner. USDA would be the data owner and MDA would be the data custodian.
- The data custodians are responsible for quality assurance and quality control and maintenance of their datasets.

2.3 Data Owners

- The authoritative data source;
- Can authorize or deny access to the data and is responsible for its accuracy, integrity, and timeliness

2.4 Open Data Council

- Provides guidance and policy recommendations and when appropriate recommend legislation and regulations for:
 - Procedures, standards, and other deliverables for open data, including for open data portals
 - Promotion, advertising, and marketing of open data
 - Best practices for sharing open data while taking into account privacy and security concerns
- Coordinates the appropriate staff at each State entity for the development, maintenance, and use of open data and open data portals
- Identifies the collective cost of operating and investing in open data and funding mechanisms to support open data; and
- Advises the Governor and General Assembly on budget matters related to open data
- Invites and encourages local entities and the legislative and judicial branches to:
 - Use open data portals established by State entities;
 - Create their own open data portals; and
 - Adopt policies consistent with the open data policy
- Establishes a plan for providing all open data to the public at no cost
- Advocates for sound records management and data preservation practices
- Makes recommendations to ensure that the purchase of new data processing devices, systems, and software by the State includes a review of compliance with the open data policy

2.5 Open Data Council Membership

Open Data Council Members positions are appointed by the governor. To apply for an open position, visit the Governor's Appointments Office website:

<https://govappointments.maryland.gov/>

After submission, the application is sent to the Department of Information Technology's Data Services Director. If approved, then the application is given to the sitting Open Data Council for final approval.

3 GENERAL DATA SPECIFICATIONS

To view instructions on adding data to the Open Data Portal refer to the Open Data Portal Tutorial for Maryland State Agencies on the following website:

<https://opendata.maryland.gov/stories/s/qr67-gp98>

3.1 Best Practices

The following are best practices from Tyler Tech, Inc.'s lesson on preparing datasets for upload to Data & Insights: <https://learn.socrata.com/prepare-your-data-for-socrata-ondemand>

Each row of a dataset should be a single instance of the "thing" being measured (for example, in a dataset measuring website visitors, it is one record of a particular site user. Or, in a dataset measuring hospital patients, it is one record of a particular patient).

Each column of a dataset should be one type of data—a characteristic that the record being measured inherently has (for example, in a dataset measuring website visitors, it might be a timestamp indicating the time and date that user visited. Or, in a dataset measuring hospital patients, it is the medical issue that patient is experiencing). Each column should have a header, or name, indicating what characteristic that column is measuring. When naming columns, avoid abbreviations and acronyms to ensure that the widest possible audience can interpret the dataset.

Do not include any empty rows or columns where feasible.

Do not include symbols (% , \$, etc.) in the cells of your dataset. Data can be configured once it is uploaded to the portal. However, decimals are acceptable, as are parentheses (for denoting a negative value).

When uploading time series data—data measuring the same object/instance at different points in time to observe a change—the unit of measurement for time should have its own column, rather than being paired with data from another characteristic/field.

Image 3-1 Time Series, Incorrect Formatting

Product ID	2014 Cost Per Unit	2014 Total Units Sold	2015 Cost Per Unit	2015 Total Units Sold	2016 Cost Per Unit	2016 Total Units Sold
8564	7.25	4523	7.40	4610	7.60	4602
3698	10.50	8064	10.60	8001	11.00	7980
1125	20.98	1023	21.56	1046	22.38	1011
4823	8.45	2365	9.00	2304	9.10	2345
3445	10.28	10200	10.99	10236	11.50	10245

Image 3-2 Time Series, Best Practices Formatting

Product ID	Year	Cost Per Unit	Total Units Sold
8564	2014	7.25	4523
8564	2015	7.40	4610
8564	2016	7.60	4602
3698	2014	10.50	8064
3698	2015	10.60	8001
3698	2016	11.00	7980
1125	2014	20.98	1023
1125	2015	21.56	1046
1125	2016	22.38	1011
4823	2014	8.45	2365
4823	2015	9.00	2304
4823	2016	9.10	2345
3445	2014	10.28	10200
3445	2015	10.99	10236
3445	2016	11.50	10245

Data & Insights can interpret columns with date and time values. When uploading a dataset with a date and time column, use one of the three following formats:

1. ISO 8601 Subset
yyyy-MM-dd
2. ISO 8601 Subset for Timezone
yyyy-MM-dd[‘T’]HH:mm:ssZ
3. American date format
MM/dd/yyyy

Data & Insights can interpret columns with location values. When uploading a dataset with a location column, use one of the following formats:

1. Exact latitude and longitude in one column
Example: (38.898303, -77.036561)
2. Latitude and longitude separated into two columns
Example: *Latitude* *Longitude*
 38.898303 -77.036561
3. American Street Address in one column (note the comma-separations)

Example: 1600 Pennsylvania Avenue SE, Washington, DC, 20003

4. American Street Address in several columns

Example:	<i>Street Address</i>	<i>City</i>	<i>State</i>	<i>Zip Code</i>
	1600 Pennsylvania Avenue	Washington	DC	20003

Prior to upload, data should be cleaned for duplicate records, erroneous or missing information, and formatting errors.

4 METADATA

Datasets are the basis for many applications and services. It is essential that the datasets be thoroughly documented to ensure the integrity of the information being presented. By properly describing what is in its corresponding dataset, metadata allows data users to make informed decisions on the suitability of data for a given purpose and to understand the method used to capture the data and its currency.

4.1 Metadata Requirements

Table 4-1 shows the required metadata fields that must be submitted with all data.

Table 4-1 Required Metadata Fields

Field	Description
Dataset Title	This is the name of the dataset. <ul style="list-style-type: none">• Give the dataset an accurate and precise title.• Spell out abbreviations and acronyms.• Do not use all caps and avoid using “data” or “dataset.”• The title should be comprehensible to the general public such as, “Minority Business Enterprise (MBE): Statewide Directory of Certified Businesses.” A bad example is, “MBE_DATA_2016”
Brief Description	Briefly describe the dataset, including the format and source of the data. This is the only large text field in the metadata, so it also serves as a catchall for other information about the dataset. This is the best place to spell out data collection, reporting, updating procedures, and workflows.
Category	Categorize the dataset from the dropdown of available categories, e.g., Public Safety, Environment, and Health. If the dataset falls into more than one category, use best judgement on which is most relevant.
Tags/Keywords	Enter tags or keywords associated with the dataset. The more tags entered, the easier it is for users to find the dataset. Keywords can include the name of the agency, topics, subject areas, and programs.
License Type	The majority of agency data is produced by state employees and should therefore be assigned a Public Domain license.
Data Provided By	Enter the names of agencies and organizations that produced the data. Consider this field mandatory despite not being shown as ‘mandatory.’
Source Link	Enter the hyperlink to the dataset or to an agency’s website. Consider this field mandatory despite not being shown as ‘mandatory.’
KML File, Shape File	If the data is geographic, and either of these files are available elsewhere, please provide a URL here.
Place Keywords	Please tag the dataset with any relevant geographic keywords. For statewide datasets this will usually be ‘Maryland.’

Update Frequency	Select from the dropdown list how often the dataset is updated. Data can be updated: <ul style="list-style-type: none"> • Monthly • Quarterly • Annually • At irregular intervals Please do not select “As Needed.” Be accurate and precise. The text field, “Other Update Frequency,” is an option to describe data update cycles in more detail. Consider this field mandatory despite not being shown as ‘mandatory.’
Time Period of Content	Enter the duration of the dataset. For example, Fiscal Year 2015 to present.

Date Metadata Written	Enter the most recent date that the dataset was updated.
Jurisdiction	If an employee of the State of Maryland, select ‘State of Maryland.’ The other options are for Maryland localities’ Open Data microsites (e.g., Frederick Open Data, Mount Rainier Stat, and several in-development).
Privacy Settings	Set the dataset for private or public usage. Note that setting it to private initially will allow the user to continue to edit the dataset until it is ready for public consumption.
Contact Email	For any questions on the dataset, enter the email contact. Consider this field mandatory despite not being shown as ‘mandatory.’

5 DATA MAINTENANCE

It is recommended that each state agency have in place an overall data maintenance plan. The plan should detail how the agency will handle its maintenance responsibilities including agency-specific QA/QC methodology and data standards.

The data owner/custodian will be responsible for ensuring the data is kept up-to-date. An automated data contemporaneity or “freshness” check is run on a daily basis to assess whether all data on the Open Data Portal has been updated in accordance with the update cycle listed in the metadata. This check is populated at a dataset on the Open Data Portal and made available as a visualization through a dashboard:

<https://opendata.maryland.gov/Administrative/Dataset-Freshness-Report-for-opendata-maryland-gov/8>

6 QUALITY ASSURANCE / QUALITY CONTROL (QA/QC)

6.1 Data Owners/Custodians

Data owners/custodians will be responsible for QA/QC of these data.

6.2 DoIT Staff

When the data is submitted for inclusion in the Open Data Portal, DoIT staff will perform some basic quality inspection prior to approving it for public availability. This includes checking for dataset naming convention and metadata completeness. Additionally, while checks for data quality and completeness is the responsibility of the data owner/custodian, DoIT does run regular automated checks to identify blank or null values in a dataset.

7 DATA SECURITY

The Open Data Portal contains the following levels of security:

- Public
- Private

8 PLATFORM SECURITY

For the safety and security of the platform and its data, roled-users are managed by DoIT Data Services staff based on their access. User account management is performed based on the following standards:

- Accounts that are inactive for more than one year are removed and any owned public or shared private content is transferred within the agency or to the Open Data Administrator account.
- New users who are invited and have not logged in within 3 months (90 days) are removed.

Only state employees are added as roled-users, but citizens can sign up for a community account that provides the same level of access as a viewer account.

9 ORGANIZATION

8.1 Organization of Data

Data will be grouped in the following categories.

Category Name	Description	Examples of data Included in Category
Administrative	Functions of businesses and organizations	Budget, Performance Improvement, Procurement
Agriculture	Rearing of Animals; Cultivation of Plants	Agriculture, Aquaculture, Herding, Irrigation, Livestock
Biota	Flora and/or Fauna in the Natural Environment	Biological Resources, Ecology, Habitat, Sea Life, Vegetation, Wildlife, Wilderness
Boundaries	Governmental Jurisdictions; Physical Delineation	County, Legislative Districts, Municipalities, School Districts, Shorelines, Voting Districts
Business and Economy	Business Activities; Economic Activities	Business, Commerce, Incentives Programs

Demographic	Census Boundaries; Characteristics of Population	Age, Blocks, Block Groups, Census Tracts, Housing, Income, Population, Urban Areas
Education	Education facilities; educational resources	K-12 schools, universities, colleges, libraries
Elevation	Height Above or Below Sea Level	Bathymetry, Digital Elevation Models (DEMs), LiDAR, Slope, TINs
Energy and Environment	Environmental Resources; Environmental Protection; Conservation Programs	Environmental Impact, Environmental Pollution, Nature Reserves, Protected Lands, Water & Air Quality, Waste Storage and Treatment

Geoscientific	Information Pertaining to Earth Sciences	Earthquakes, Erosion, Geology, Geophysical Features, Hydrogeology, Minerals, Soils
Health and Human Services	Health Facilities; Health Services; Health Trends; Human Ecology	Diseases, Health Care Facilities, Mental and Physical Health, Public Health, Substance Abuse
Historic	Archeological Data; Historic Locations	Archeological Sites; Historic Properties; Preservation Sites; Sites of Historic Events
Housing	Housing-related data such as financial information and sales data.	Maryland Total Residential Sales; Maryland Acres for Residential Development; loan programs
Hydrology	Movement of Water On and Below Earth's Surface and In the Atmosphere	Continental Shelf, Currents, Dams, Hydrography, Outer Continental Shelf, Rivers, Streams, Tides, Shoreline, Watersheds, Wetlands
Imagery	Imagery	Color Infrared Imagery; Composite Imagery; Grids; High Resolution Imagery; Incident-Specific Imagery
Location	Positional Information and Services	Addresses, Centroids, Geodetic Control Points

Military	Military Activities, Bases and Structures	Barracks, Events, Facilities, Military Transportation, Training Grounds
Planning	Land Designations; Property Boundary Delineation; Tax Appropriations	Cadastral Surveys, Easements, Land Cover, Land Use, Parcel Boundaries, Tax Maps, Zoning
Public Safety	Prevention of and Protection From Crime, Damage, Danger, Harm or Injury	Correctional Facilities, Crime Statistics, Emergency Response, Fire and Police Stations
Society	Community and Cultural Facilities; Non-Demographic Characteristics	Community Centers, Cultural Resources, Libraries, Monument, Schools, Sporting Venues

Structure	Man-Made Construction With No Cultural Value	Architectural and Structural Plans, Building Footprints, Dams, Towers
Transportation	Conveyance of Persons and/or Goods; Mobility Resources; Modal Infrastructure	Airports, Bridges, Nautical Charts, Roads, Railways, Shipping Lanes, Tunnels
Utility	Communication Systems; Energy Systems; Waste Systems; Water Systems	Broadband, Cellular, Electric, Fiber, Gas, Geothermal, Hydroelectric, Radio, Sewage, Solar, Water
Weather	Atmospheric Conditions; Atmospheric Phenomena	Atmospheric Conditions; Atmospheric Phenomena

10 PROCEDURES FOR PUBLICATION OF DATA

For procedures on adding data to the Open Data Portal refer to the Open Data Tutorial for Maryland State Agencies on the following asset:

<https://opendata.maryland.gov/stories/s/qr67-gp98>