

**OPEN DATA POLICY REQUIREMENTS**

**5160.1**

(Revised 03/2019)

Effective July 1<sup>st</sup>, 2019, as part of the Open Data policy, each Agency/state entity shall:

1. Build or modernize Information Technology (IT) solutions in a way that maximizes interoperability and information accessibility. Although this policy does not require Agency/state entities to modernize existing IT solutions, it does require data considerations identified in this section be applied when a state entity undertakes a modernization effort that substantially modifies an existing IT solution.
  - a. Exercise forethought when architecting, building, or substantially modifying an IT system to facilitate data distribution to the public, where appropriate.
  - b. Use machine-readable and open formats for information as it is collected or created. Where applicable, machine-readable and open formats must be used in conjunction with electronic or paper-based information collection efforts.
  - c. Prioritize the use of open formats that are non-proprietary, publicly available, and that place no restrictions upon their use.
  - d. Apply open licenses, such as Creative Commons Zero (CC0), to information as it is collected or created so that if data are made public there are no restrictions on copying, publishing, distributing, transmitting, and adapting.
  - e. Systems must be scalable, flexible, and facilitate extraction of data in multiple formats and for a range of uses as internal and external needs change, including potential uses not accounted for in the original design (e.g. leveraging standards and industry best practices for information sharing, separation of data from the application layer to maximize data reuse opportunities.)
2. Whenever feasible, make data broadly available to the public through the Agency/state entity's open data site or portal, pursuant to the limited exceptions outlined in SAM Section 5160.2.
3. Describe information using standard metadata as the data is collected or created.
  - a. Open Data shall include Project Open Data Catalog Vocabulary (DCAT) standards modified for California, see [Open Data Handbook](#) for specifications and formats.
  - b. Agencies/state entities may expand upon metadata and data dictionaries based on standards, specifications, or formats developed within different communities (e.g., financial, health, geospatial, law enforcement). Groups

(Continued)

**OPEN DATA POLICY REQUIREMENTS**

**5160.1 (Cont. 1)**

(Revised 03/2019)

that develop and promulgate these metadata specifications must review them for compliance with DCAT specifications and formats.

- c. Metadata and data dictionaries shall be in a machine-readable format to provide users the ability to export when needed.
4. Adopt effective governance and data asset portfolio management approaches, including data management and release practices to ensure consistency.
  - a. Create and maintain an Agency/state entity enterprise data inventory, see [Open Data Handbook](#) for inventory specifications and formats.
  - b. The inventory shall indicate, as appropriate, if the Agency/state entity has determined that the individual datasets may be made publicly available (i.e., release is permitted by law, subject to all privacy, confidentiality, security, Agency/state entity has ownership of data, and other valid requirements) and whether they are currently available to the public.
  - c. The inventory shall list any datasets that can be made publicly available at the Agency/state entity's open data site or portal in a format that enables automatic aggregation by Data.ca.gov and other services (known as "harvestable files"), to the extent practicable. See [Open Data Handbook](#) for best practices, tools, and schema to implement the public data listing and harvestable files.
  - d. Public data listing should include, to the extent permitted by law and existing terms and conditions, datasets that were produced as a result of legislative mandates, state grants, contracts, and cooperative agreements (excluding any data submitted primarily for the purpose of contract monitoring and administration), and, where feasible, be accompanied by standard citation information, preferably in the form of a persistent identifier.
  - e. Assign a Data Coordinator to coordinate and maintain Agency/state entity's public data. The Data Coordinator's contact information must be identified in the enterprise data inventory.
5. Prioritize the collection of data sets.
  - a. Agencies/state entities shall identify and engage stakeholders as part of the intake process.
  - b. Create a process to engage with customers to solicit help in identifying data sets of value to the public, in prioritizing the release of public datasets

(Continued)

**OPEN DATA POLICY REQUIREMENTS**

**5160.1 (Cont. 2)**

(Revised 03/2019)

and determining the most usable and appropriate formats for release. Agencies/state entities should make public data available in multiple file formats according to their customer needs (e.g. high-volume datasets of interest to developers should be released using bulk downloads as well as Application Programming Interfaces (APIs)).

6. Ensure that privacy and confidentiality are fully protected, and that data is properly secured.
  - a. Leverage an internal data governance process to determine if information collected or created can be made publicly available or is subject to restrictions (e.g. privacy, confidentiality, security, trade secret, contractual). See [Open Data Handbook](#) for additional information.
  - b. If the Agency/state entity determines that information should not be made publicly available on one of these grounds, the Agency/state entity must document this determination through its internal data governance process.
  - c. Consider security-related restrictions including National Institute of Standards and Technology (NIST) Federal Information Processing Standard (FIPS) Publication 199 “Standards for Security Categorization of Federal Information and Information Systems,” which includes guidance and definitions for confidentiality, integrity, and availability.
  - d. Collect or create only that information necessary for the proper performance and evaluation of Agency/state entity functions and which has practical utility. Limit the collection or creation of information which identifies individuals to that which is legally authorized and necessary for the proper performance of Agency/state entity functions.
  - e. Limit the sharing of information that identifies individuals or contains proprietary information to that which is legally authorized and impose appropriate conditions on use where a continuing obligation to ensure the confidentiality of the information exists. [Data sharing agreements](#) must be created to exchange information across Agencies/state entities and with research institutions in compliance with the State’s information security and privacy policy and standards, see SAM Section 5300 and Statewide Information Management Manual (SIMM) Section 5305-A.
  - f. Ensure that information is protected commensurate with the risk and magnitude of the harm that would result from the loss, misuse, or unauthorized access to or modification of such information. Agencies/state entities shall consider the standard for information classification detailed in SIMM Section 5305-A and other publicly available information when

SAM – INFORMATION TECHNOLOGY  
(California Department of Technology)

Continued)

**OPEN DATA POLICY REQUIREMENTS**

**5160.1 (Cont. 3)**

(Revised 03/2019)

determining whether information should be considered Personally Identifiable Information.