



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

**TITLE:** Open Data Portal Policy  
**NUMBER:** BUL-065104  
**ISSUER:** Oscar Lafarga, Executive Director  
Office of Data and Accountability

## ROUTING

All Employees  
All Locations

**DATE:** March 8, 2019

**POLICY:** The purpose of this bulletin is to define the policy of the LAUSD Open Data Portal. The LAUSD Open Data Portal is a website that provides the public access to District data through reports and downloadable, machine-readable datasets.

**BACKGROUND:** As stated in board resolution [Res-021-17/18](#), the Los Angeles Unified School District is committed to openness and transparency based on the belief that open government improves democratic functioning and fuels innovation and collective problem solving. To deepen this commitment, the District has created an Open Data Portal, where all datasets shall be publicly-accessible to the extent permitted by law and subject to valid privacy, confidentiality, security and other legal or practical restrictions.

Although the Office of Data and Accountability (ODA) manages Open Data, this initiative is a collaborative effort involving other divisions that contribute operational and performance data to the portal. Examples of data include, but are not limited to, the following: results on annual high-stakes assessments, graduation rates, attendance, enrollment, college readiness data, budget and finance information and facilities data. In addition, the Informational Technology Division (ITD) provides the necessary technological resources and support to make the data easily accessible to the public.

## I. GOALS AND OBJECTIVES

The primary goals of the LAUSD Open Data Portal are as follows:

1. Build trust and accountability in the District by shedding light on District practices through data.
2. Empower stakeholders to participate in District oversight through increased access to data.
3. Create a central site where parents, school communities, researchers, advocacy groups, and the community-at-large can access school and District-level data.
4. Make datasets available to the extent permitted by law and subject to valid privacy, confidentiality, security and other legal or practical restrictions.



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

5. Experiment with new technologies, visualizations, and applications with the goal of identifying cost-effective solutions to improve our schools.

## **II. ROLES AND RESPONSIBILITIES**

### **A. EXECUTIVE DIRECTOR, OFFICE OF DATA AND ACCOUNTABILITY**

The Executive Director of the Office of Data and Accountability provides the overall governance of the LAUSD Open Data Portal.

### **B. OPEN DATA OFFICER**

The Open Data Officer, working under the guidance and supervision of ODA's Executive Director, coordinates the release of datasets, delivers support to data-sharing divisions and provides assistance and outreach to end-users accessing the data. The Open Data Officer also works closely with community members to heighten awareness of, and improve access to, the data.

The Open Data Officer plans, organizes, coordinates, and implements all aspects of the District Open Data program by:

1. Ensuring the Open Data Portal is operational and includes the most recent datasets available.
2. Managing and coordinating data for release on the Open Data Dashboard and in the Open Data Catalog.
3. Guiding the District's cross-divisional effort of collecting data sets and preparing them for public view.
4. Ensuring that data is available, reliable, consistent, accessible, secure, and timely to support the mission and activities of the District.
5. Collaborating with other divisions and the public to promote the Open Data portal.

### **C. OPEN DATA DESIGNEE**

Each division contributing datasets to the Open Data portal will select at least one designee who will be responsible for:

1. Working with the Open Data officer and their division head to identify data sets to publish.
2. Uploading datasets to the Open Data Catalog through the eLibrary file upload process (described below).
3. Ensuring that all datasets meet the recommended standards described in Attachment A.
4. Providing support and assistance to users accessing their Division's datasets.

### **D. OPEN DATA WORKING GROUP**

Open Data designees participate in the Open Data working group. The group:



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

1. Ensures that all appropriate datasets are identified for public disclosure.
2. Meets regularly to discuss, plan, and prepare the release of datasets to the public.
3. Ensures that Open Data standards are applied to datasets published on the portal.
4. Provides feedback on the functionality of the Open Data Portal to ensure public access to data are optimized.

## III. OPEN DATA PORTAL

The LAUSD Open Data Portal serves as the central source for open data, where District datasets shall be made available to the extent permitted by law. The portal comprises two components for viewing and accessing data: a dashboard display of data and a catalog of downloadable and machine-readable datasets.

### A. OPEN DATA DASHBOARD

The Open Data Dashboard displays a cross-section of key performance indicators, many of which are represented on the District's Local Control and Accountability Plan. The data are displayed in graphical and tabular format and allow for comparisons over time, across Local Districts and Board Districts, by school and for specific student groups. The dashboard also includes school-level profiles, so that schools can share relevant data with parents and other stakeholders in a printer-friendly format. New metrics will be added to the dashboard over time, but currently the following types of data are available:

1. Student Demographics
2. Attendance
3. Student Outcomes
4. College and Career Readiness
5. College Persistence

### B. OPEN DATA CATALOG

The Open Data Catalog houses datasets containing school performance and school operational data. The catalog includes datasets of the indicators presented on the Open Data Dashboard along with other non-dashboard datasets. For the catalog's initial release, the District has prioritized student performance, budget and finance, and facilities data. Over time, the catalog will continue to grow to include datasets from divisions across the District.

All datasets in the Open Data Catalog share the following characteristics:

1. Datasets are shared in a secure fashion consistent with the federal Family Educational Rights and Privacy Act (FERPA) and other applicable laws to ensure that student privacy is not compromised.
2. Datasets are centralized and easy to locate



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

3. Datasets are downloadable, machine-readable and processable, and, where appropriate, provided in a non-static format.
4. Datasets are searchable using common sense search terminology.

## IV. DATA MANAGEMENT

The Open Data Portal sources data from District transactional and reporting systems and from publicly-accessible web pages, both District-managed and external, such as the California Department of Education. To ensure consistency in the presentation, formatting and storage of data in the data catalog, this policy outlines data standards and the process for divisions to make their data available through the eLibrary file management system.

### A. DATA STANDARDS (ATTACHMENT A)

Adhering to a set of data standards will ensure that all data residing on the Open Data Portal are easy for the public to access, understand and make use of. Data standards ensure that data are presented in a consistent format regardless of the contributing organization or division. The complete list of standards can be found in Attachment A and include guidance related to: appropriate file types, the use of unique identifiers, data dictionaries, data type formats, suppression rules to protect student privacy and tagging datasets with pertinent metadata.

### B. SHARING DATA THROUGH THE CATALOG (ATTACHMENT B)

The following steps describe how data-contributing divisions will share data through the catalog and how the datasets are approved for release:

1. Each division's Data Designee will prepare their datasets to upload through the eLibrary system identifying the dataset as "open data".
2. The Open Data Officer will review and approve these datasets based on the standards described in the standards document
3. Open Data datasets will be set aside in the eLibrary system to be accessed and reviewed.
4. The Office of the Superintendent eLibrary designee will review the Open Data datasets for content and provide final approval to uploaded and be made available for public access.

### C. TRAINING FOR OPEN DATA DESIGNEES

ODA and ITD have developed a training curriculum available on MyPLN, which takes new open data contributors through a step-by-step process for how to contribute data to the portal. The training describes the recommended data standards and teaches contributors how to upload datasets to the Open Data catalog.



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

## D. DATA PRIVACY

The Dashboard does not allow users to drill to student-level detail. To protect student privacy, data are not reported if there are fewer than 11 students in a group or at a school site. These rules apply to both dashboard reports and data catalog datasets of student information. These rules were established to protect student privacy as governed by the Family Educational Rights and Privacy Act (FERPA).

### **RELATED RESOURCES:**

*CDE Data Quest:* [https://dq.cde.ca.gov/dataquest/State Accountability Report](https://dq.cde.ca.gov/dataquest/State%20Accountability%20Report)  
*Card:* <https://www.cde.ca.gov/ta/ac/sc/>

*California Assessment of Student Performance and Progress (CAASPP):*  
<https://caaspp.cde.ca.gov/>.

### **ATTACHMENTS:**

Open Data Portal Data Sets Standards (Attachment A)  
Open Data Portal Process Flow (Attachment B)

### **ASSISSTANCE:**

For assistance with the Open Data Portal Dashboard, contact the Open Data Officer from the Office of Data and Accountability at (213) 241-2460.

For eLibrary assistance, contact ITD at (213) 241-5200.



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

## Attachment A

### Open Data

#### Dataset Standards

This document describes standards for datasets uploaded to the Open Data catalog. The following standards not only make it easier for the public to access, understand, and make use of datasets that are made available through Open Data but also ensure that information is presented in a consistent format regardless of the contributing organization:

- **Upload datasets in a machine-operable format.**
  - This format meets the requirements of an open format that can be accessed by proprietary or non-proprietary software.
  - The preferred format is .csv, but if this is not possible then an .xlsx format is acceptable.
  - To allow for machine-operability of the data, PDFs are discouraged for upload into the Open Data Portal.
- **Include column headers with each column of data.**
  - Remove spaces from column headers and use camel casing (e.g., CostCenterCode)
  - Do not include special characters such as % or # symbols.
  - Use column headers that are succinct but easy to understand (e.g., Enrolled rather than Number Enrolled).
- **Within each dataset, include all locations that are associated with the data being represented.**
  - This approach will not only benefit public users by providing a comprehensive, singular data set for them to download and analyze (e.g., Smarter Balanced Assessment results dataset for all participating schools and grade levels), but will also eliminate the need for users to download and merge data sets in order to compare two or more schools.
- **Include a standard set of columns, in the following order, with each data set:**
  - Cost Center Code (displayed as CostCenterCode) – The inclusion of a unique identifier will allow data consumers to associate dt associated with the same locations.
  - Location Code (LocationCode) – The unique 4-digit code associated with the location (not the preferred location code, also known as Campus Code)
  - CDE CDS Code (CDSCode) – The unique 14-digit code assigned to each California public school by the California Department of Education.
  - School Name (School) – Long School Name is preferred (e.g. Phineas Banning Senior High School rather than Banning SH)
  - School Year or Data Date – See next entry.
- **Include a data date to specify**  
the timeframe represented by the data (Note: this will vary by dataset).



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

## Attachment A

- Date formats may include but are not limited to:
  - YYYY-MM-DD
  - YYYY-MM
  - YYYY-YYYY
  - Other: Some datasets, for example, assessment results, may make reference to an assessment window instead; e.g., '2017-2018 Beginning of Year'
- Column headers will vary depending on the date being represented; 'DataDate' is acceptable, though 'SchoolYear' is preferable when the data represent the entire school year.
- **Other location attributes will be made available to data consumers through a master location file:**
  - The Office of Data and Accountability will upload a master location file to the eLibrary, so that designated contributors do not have to add the same set of location attributes (local district, board district, etc.) each time.
  - A separate master location file will exist for each school year.
- **Standardize indicators associated with school performance data**
  - Display total counts of students in the population (e.g., tested, enrolled, etc.), counts of students within each performance area/level and percent of students within each area;
  - Display percentages as a number between 0 and 1 rounded to 4 decimal places (e.g., 10.43% would be stored as .1043 in the file)
- **For student performance data, include rows for District-level summary**
  - Include school ('Los Angeles Unified School District'), cost center code (0000000), location code (0000), CDE CDS Code (19647330000000)
  - This information will be useful in the event of school-level results being suppressed (small n sizes). To find out more, see below.
- **For student performance data, only include data for one school year**
  - Since location information may change from one school year to the next, a file containing only one school year, will allow data consumers to associate the dataset to the relevant location master file.
- **Apply suppression rules on aggregated student data that comply with FERPA guidelines.**
  - To ensure confidentiality of individual student results, State guidelines for minimum cell size for reporting data stipulate that if the number of students included in calculations is less than 11 then data must be suppressed.



# LOS ANGELES UNIFIED SCHOOL DISTRICT POLICY BULLETIN

---

## Attachment A

- **Use a standard file naming convention for documents uploaded to the Open Data catalog.**
  - Keep filenames short. Though the technical limit is 255 characters, a good rule of thumb would be 50 characters or less.
  - Apply the following file naming convention standards: no special characters, no spaces, do not begin a file name with a number, use camel-casing, apply date format or school year suffix to the end of the file (e.g., '20172018')
  - Example: SmarterBalanced20172018
- **Include a data dictionary with each dataset being uploaded to Open Data.**
  - The data dictionary will be added as a .csv or excel file attachment to the primary data set. [Click here for instructions on how to add an attachment.](#)
  - The filename should include the name of the accompanying dataset with 'DataDictionary' appended to the end.
  - At a minimum the following fields should be included in each data dictionary:
    - Column Name – As it is represented in the data set
    - Description – brief description of the information contained in the field
    - Data Type – String/Text, Numeric
  - Two data sets pulled from the same source system should refer to the same information using the same column headers (e.g., CostCenterCode vs. CostCenter).
  - Two data sets pulled from different source systems may not necessarily refer to the same information using the same column headers (e.g., Fund Center vs. Cost Center). In these cases, we will use the datasets's data dictionary or the glossary of terms to explain these differences.
- **Include metadata details through the eLibrary upload process.**
  - Topic – Select the topic that applies to the dataset being uploaded
  - Title – Start the document title name with subject matter associated with the dataset, so that similar datasets will appear together in sorted catalog list (e.g., Smarter Balanced 2017-2018 rather than 2017-2018 Smarter balanced)
  - Summary – Include the following elements:
    - A brief description of the dataset;
    - What data are contained in the dataset (e.g., # of students, % of students meeting standards, etc.)
    - Data source (e.g., MiSiS, CDE, CALPADS, etc.)
    - Data timeframe (e.g., month-to-date, year-to-date, as of YYYY-MM-DD, Annual, etc.)
    - Data refresh (e.g., Weekly, Monthly, Annually, etc.)



## OPEN DATA PORTAL PROCESS FLOW

