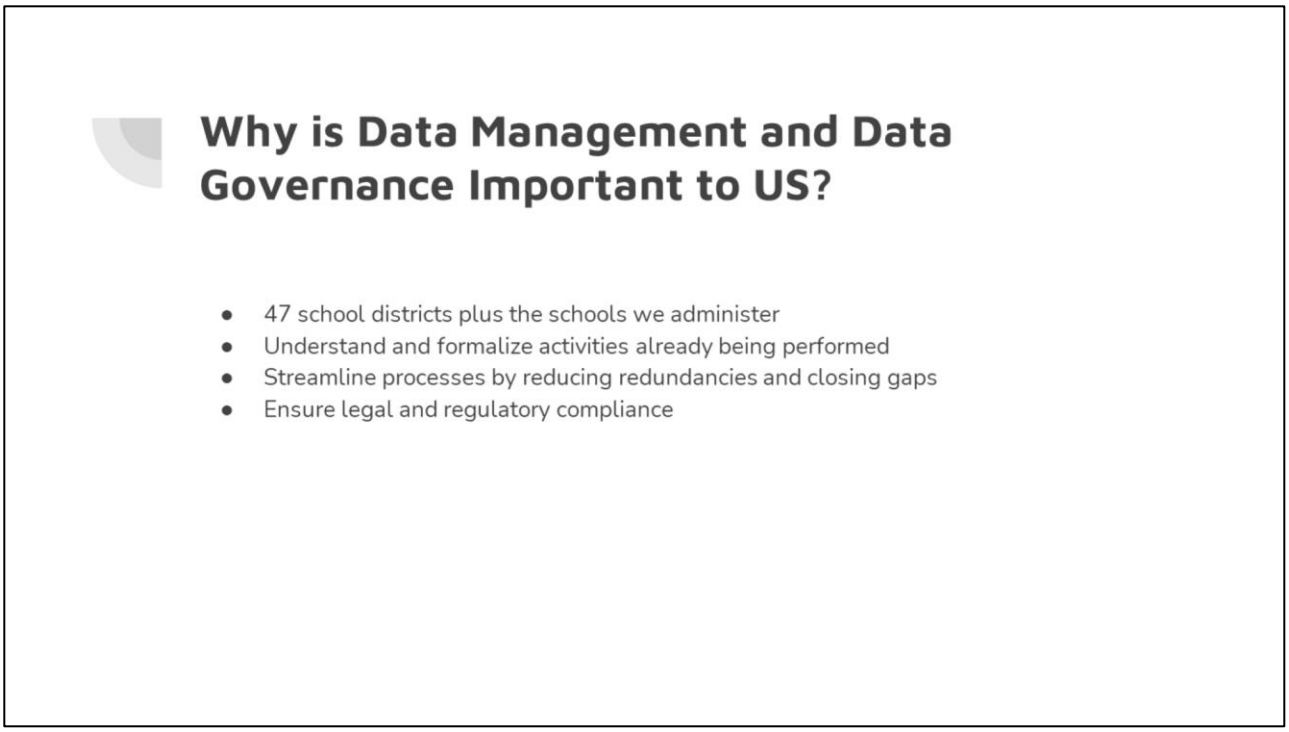
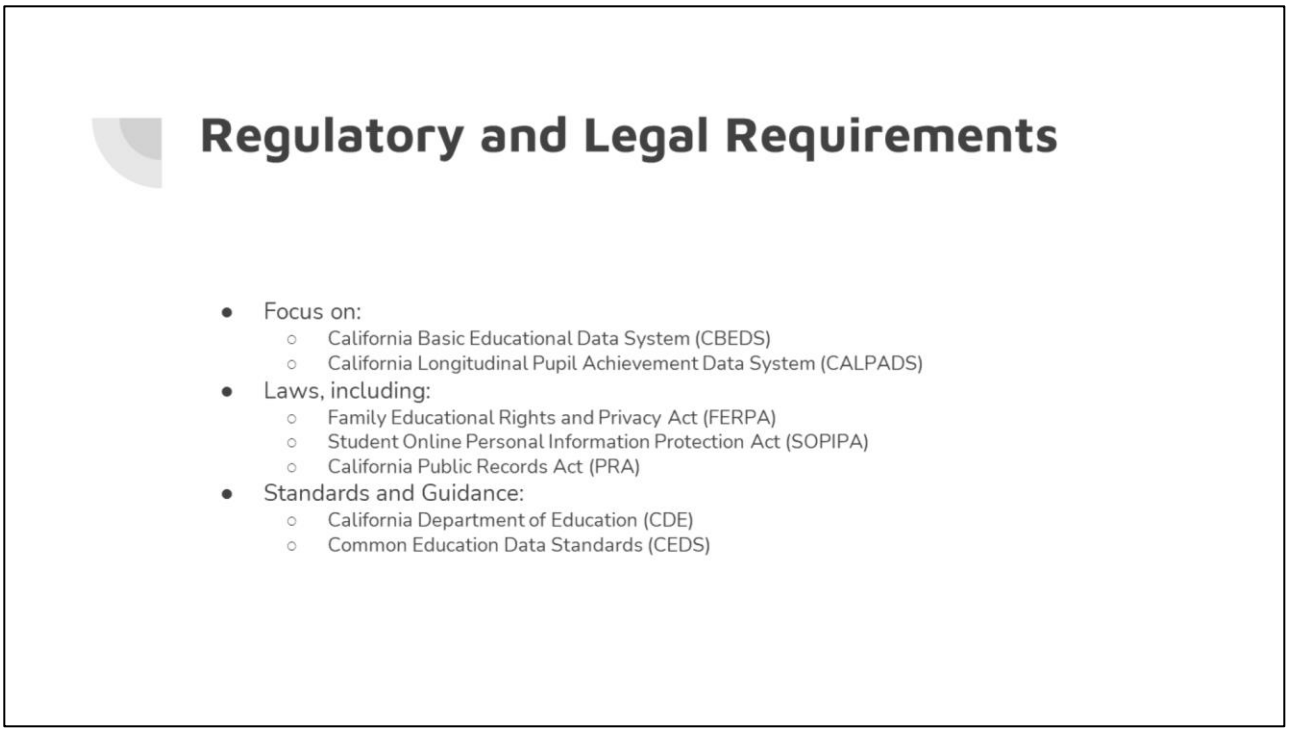
Data management provides value by providing a laser focus on data and how it is handled. Data governance makes sure that data management is organized and not ad-hoc; governance activities are the rule making or framework implementations that ensure data management is consistent. One framework is for data management to look at how data is created, stored, accessed, processed, and eventually destroyed. One important thing to note is that data management is not the same as IT or database administrator; though they could, and probably should, form part of the data governance team. One reason we refer to the data management team as a data governance team is to differentiate

between the technical IT activities and the management of data as an asset (and liability). Data governance is defining activities and oversight of those activities, data management is the execution of the defined activities.



Data management and governance is important to the Kern County Superintendent of Schools (KCSOS) because we already handle massive amounts of data and are required to report data to many different entities. Along with the schools we administer, we also serve 47 other districts within the county. By formalizing data governance, we will be able to address redundancies and gaps while ensuring regulatory and legal compliance. One thing to really consider is that the KCSOS and the 47 districts we serve already gather and disseminate data. Creating a data governance program will not shake things up, it will just formalize what is already being done while remedying and gaps that may be

discovered.



Education rightly has many stakeholders and is highly regulated. There are many legal and regulatory requirements that touch upon education including Family Educational Rights and Privacy Act (FERPA), Student Online Personal Information Protection Act (SOPIPA), and California Public Records Act (PRA). For relative simplicity I would like to focus on 2 regulatory requirements for KCSOS and the school districts we serve. All are required to report data to both the California Basic Educational Data System (CBEDS) and the California Longitudinal Pupil Achievement Data System (CALPADS). The California Department of Education (CDE) data standards are guided by the

Common Education Data Standards (CEDS) provided by the National Center for Education Statistics (NCES).

Plan

- Gather Information
- Define Stakeholders, Processes, Frameworks
- Define Data

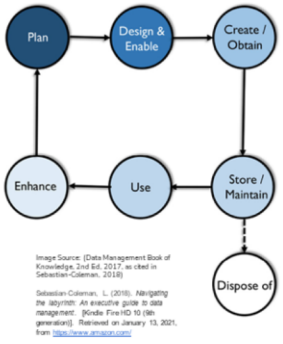


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

- Plan. The first step is non disruptive. This proposal and presentation are examples of the first step. It is information gathering and just knowing which processes are in place right now. We already have many different data reporting requirements and many different data reporters. We just want to get a complete picture of who gathers and reports data and how. While we want to address all data practices and requirements, we will initially focus on CBEDS and CALPADS requirements and current practices around those requirements. But we do want to have a complete understanding of our data landscape

besides the two mentioned data systems we will initially focus on. With a focus on CBEDS and CALPADS requirements we can define data requirements and connect them to existing practices in order to design our data management program.

- The data lifecycle framework we present here consists of the following elements: plan, design & enable, create/obtain, store/maintain, use, enhance, and dispose of

Design and Enable

- Map what is already being done to definitions
- Design for and fill in any gaps in practice
- Train users and deploy system
- Specs to Plan phase definitions

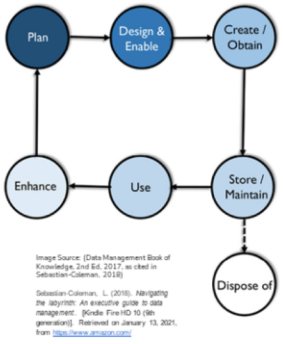


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

The design and enable phase builds upon the definitions identified in the plan phase. Here we design the infrastructure and map already existing processes to CBEDS and CALPADS requirements. All the school districts are already required to collect and provide data. The design phase will just map what the schools and districts already collect and ensure the correct granularity is provided to each data system as required. This will also ensure that only those that are required to have access to certain information are given access to the data they should have access to.

Create/Obtain

- Quality stage
- Ensure data and process quality
- Ingest and proper permissions to obtain data

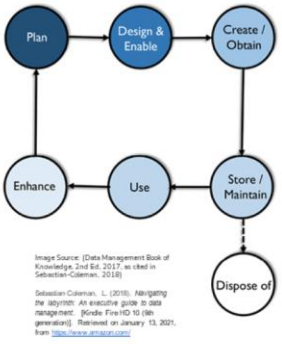


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

The create/obtain phase is where processes and people can ingest and access the relevant data, they have access to. This phase is where you ensure data quality and ingestion and reporting quality. In short, this phase ensures that the quality is where you want it to be. While the individual schools will have access to much more data about students and employees, CBEDS and CALPADS have different levels of access and granularity. This phase examines and ensures that is the case.

Store/Maintain

- Ensure data quality after ingestion
- Migrate data to proper data systems
- CBEDS and CALPADS
- Proper data granularity
- Analyze internal quality issues
- Analyze external quality issues

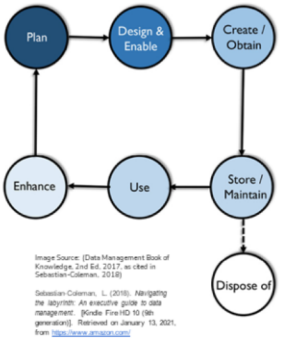


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

The store/maintain phase is always important, but it is extremely important to us right now. Some of the data that previously needed to be reported to CBEDS now needs to be reported to CALPADS; and while reporting to CALPADS is required, it is still under development. There are plenty of issues with the developing CALPADS systems, there are many tickets that are being always processed. On one hand we need to ensure our processes are clean and our data quality is intact. Because of the issues with migrating data to CALPADS from CBEDS and the developing pains of CALPADS, we need to know why some quality issues arise. Is it because of local faults or is it because of the

developing and migrating issues at the state level?

Use

- Feedback
- Extract value

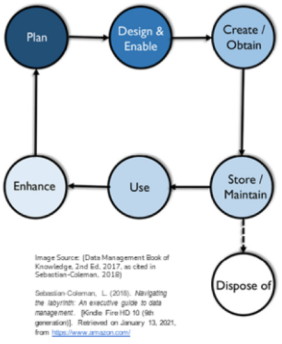


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

The use phase learns and extracts value from data and data practices. This phase also requires feedback in order to ensure practices, data quality, and access are correct, need to be changed, or improved. The use phase is pretty much after definitions and practices have been used and tested. Are we able to maximize the value and streamline the processes and standards we have in place around data? We can always improve, especially while CALPADS is being developed.

Enhance

- Enhance based on feedback
- Enhance based on changing regulations
- Enhance by addressing different stakeholders
- Enhance by evaluating the whole program

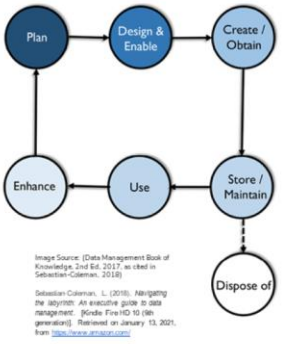


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

The enhance phase acts on the feedback from all phases, especially the use phase. Enhance also responds to arising external requirements, such as new regulations or security threats. While this initial focus is on CBEDS and CALPADS, there are many different stakeholders and regulatory bodies that require reports. While all schools and districts already collect all sorts of data, data management and data governance is required to enhance the practices. We can cut down on redundancies and gaps by having a plan, we can streamline processes to collect and protect data, we can keep up with the various external data reporting requirements. The data governance team can handle and process new

requirements and keep all the data stewards up to date without having to have them sift through all the never-ending external requirements that spring up.

Dispose of

- Limit Liability
- Archive Data
- Aggregate and Anonymize Data
- Help streamline current usage

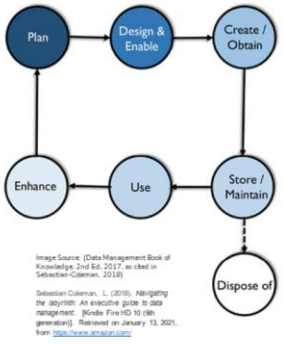
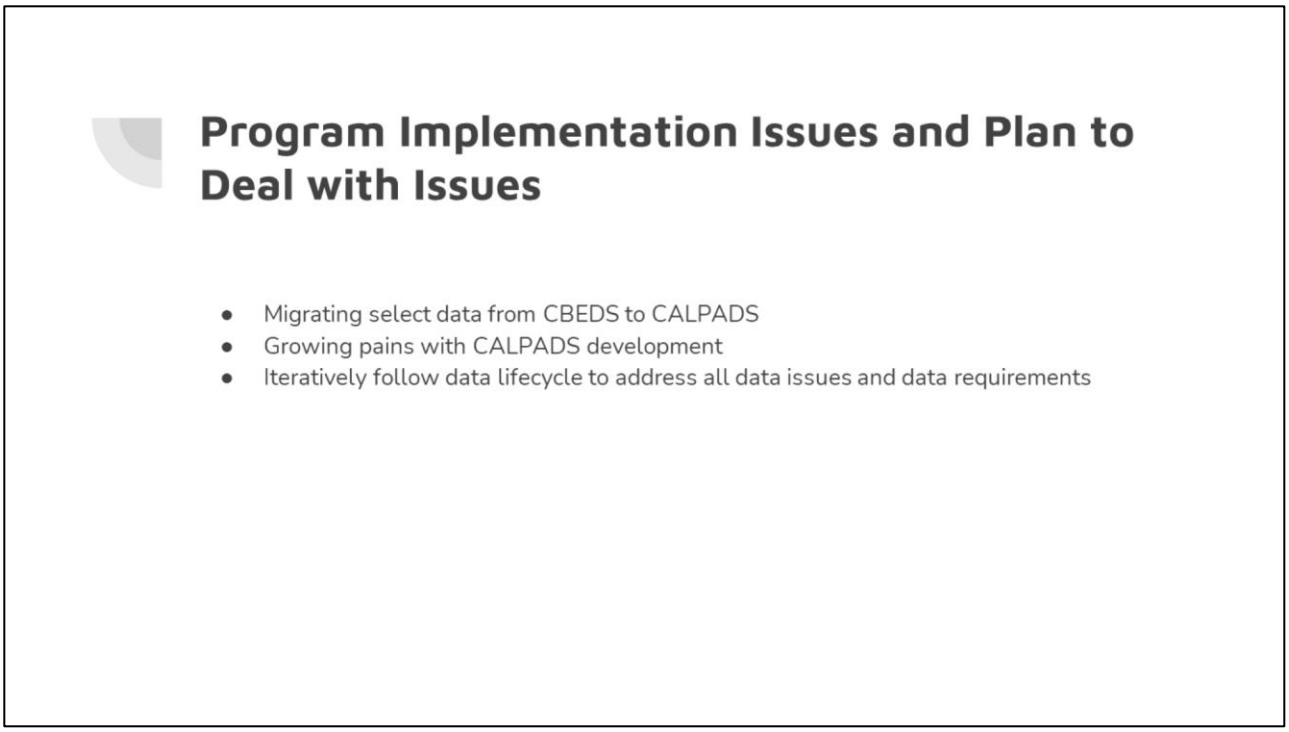


Image Source: (Data Management Book of Knowledge, 2nd Ed. 2017, as cited in Sebastian-Coleman, 2018)

Sebastian-Coleman, L. (2018). *Navigating the labyrinth: An executive guide to data management*. [Kindle Fire HD 10 (9th generation)]. Retrieved on January 13, 2021, from https://www.amazon.com/

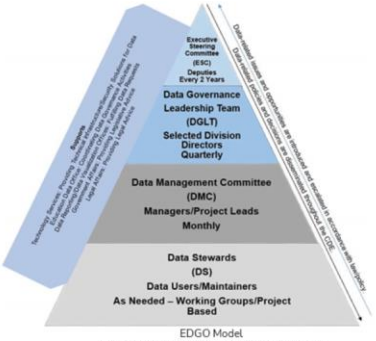
According to the model we are following the ‘dispose of’ phase branches off of the store/maintain phase, so does the use phase. For liability and storage capacity reasons we do not want to store/maintain all data indefinitely. During this phase we dispose of data or archive data that is not needed in day-to-day use. It is not arbitrary; we need to have a plan and schedule. The data governance team evaluates the legal obligations and plans accordingly. Legal obligations not only consider how long to keep records, but also court orders to keep specific records longer. The ‘dispose of’ phase needs to consider each of these issues.



One issue that will inevitably arise are the different data and granularity requirements to different data systems. Some data reporting requirements are being migrated from CBEDS to CALPADS, but CALPADS is currently being developed and reported to. That will be an ongoing issue. Also, there are many tickets that of issues with CALPADS that the state acknowledges and lists, this too will be an ongoing issue. We will have to keep up to date, and preferably be involved with the development of CALPADS by providing input to the CDE.

Long Term Maintenance Plan

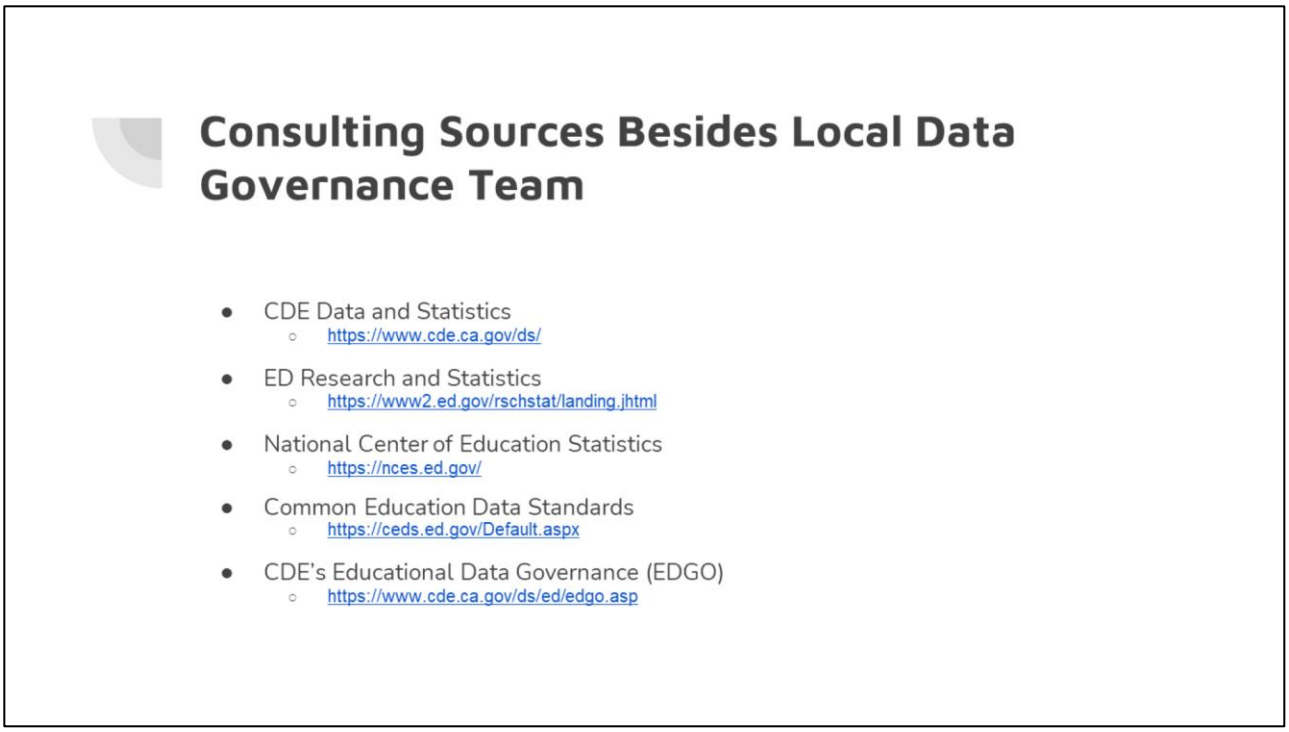
- Framework
  - CDE's Educational Data Governance (EDGO) Model
- Kern County Superintendent of Schools (KCSOS)
  - Data Governance = Superintendent Level
  - Data Management = District Level
  - Data Stewards = All Levels, schools, districts, etc.



EDGO Model
Source: https://www.cde.ca.gov/ds/ed/edgo.asp

Long term a good model would be to follow the CDE's Educational Data Governance (EDGO) program. We would have a data governance team at the superintendent level, data managers at the district level, and all the different data stewards and users at various levels. The EDGO model has a steering committee meeting every 2 years; in reality, the data governance team would probably meet with the superintendent, board of supervisors, and different district school boards much more frequently as needed and requested. But the EDGO model is a good basis. The data governance team would meet at least quarterly, the data managers at least once monthly, and data stewards much more

frequently as needed. In reality, the data governance team would be ready and meeting all the time with the different stakeholders.



Some sources to consult for more information. I would primarily focus on the California Department of Education's (CDE) Data and Statistics link: https://www.cde.ca.gov/ds/. This site has most of the information we are obligated to comply with. The U.S. Department of Education's (ED) Research and Statistics link (https://www2.ed.gov/rschstat/landing.jhtml) is also a good resource around data, data practices, and data requirements. Both the CDE and ED refer to the National Center of Education Statistics (https://nces.ed.gov/) and Common Education Data Standards (https://ceds.ed.gov/Default.aspx) they developed and update on a regular basis. I would

also recommend CDE’s link specifically on Educational Data Governance (https://www.cde.ca.gov/ds/ed/edgo.asp).