CALIFORNIA Data System

Scope and Decision-Making Frameworks for the Cradle-to-Career Data System

May 2020 Cradle-to-Career Workgroup
Kathy Reeves Bracco and Kathy Booth, WestEd

Over the first quarter of 2020, planning for the state data system outlined in the Cradle-to-Career Data System Act[1] (Act) focused on the products it would produce in the first phase of development. At the April workgroup meeting, the partner entities agreed to the following:

- Information on the progression of Californians from early years (including both education and care experiences) to K–12 to postsecondary education (including access to financial aid) to employment, made available to the public through a dashboard and a query builder tool
- The capacity to link information between partner entities to produce approved research studies and purpose-specific interagency data sets
- An online library for research conducted using linked data from the state system, along with information on data definitions
- Tools that help to streamline planning for and access to college, with the option for students to share information on their participation in social service programs in order to qualify for supports while in college

With the products established, the Cradle-to-Career Workgroup (workgroup) will now address questions of governance, which is essentially the framework for managing the state data system. The federal Statewide Longitudinal Data System Support Team provides a useful definition of data governance, which illustrates the expansiveness of the issue.

[1] Read the Act at: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=EDC&division=1.&title=1.&part=7.&chapter=8.5.&article=

Data governance refers to the overall management of data, including its availability, usability, integrity, quality, and security. It is the means by which organizations (or groups of organizations) make collaborative decisions about their collective information assets. Data governance is both an organizational process and a structure. It establishes responsibility for data, and it organizes program area and agency staff to collaboratively and continuously improve data quality and usability and to ensure data security and confidentiality through the systematic creation and enforcement of policies and procedures. Data governance also ensures efficient and effective collaboration, authority, and accountability by defining and assigning clear roles and responsibilities.[2]

Most state data systems address their governance approach through written documentation, which may be enshrined in a governance charter, a "book of data governance," or a set of bylaws. While the structure of the documents differs somewhat, each addresses a set of common components (see the box below). This paper addresses three of these topics: mission and vision statements, purpose and scope statements, and the decision-making model. The remaining topics will be addressed in the June background paper. For each component, we highlight examples of how other states have addressed the issue, explain the current status of the issue in California, describe considerations for the workgroup, and present a set of action items and questions for the May 2020 meeting.

Common Components for State Data System Governance

- Mission and vision statements
- Purpose and scope statements, often in the form of a research agenda
- Decision making and staffing model
- Internal technical and legal framework
- External data sharing policies and procedures

[2] SLDS Issue Brief, *Communicating the Value of Data Governance* https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=28771

Mission and Vision Statements

In a governance framework, vision statements articulate the problem the state data system sets out to solve and mission statements describe how the system helps to reach those goals. Mission and vision statements help with evaluating opportunities and priorities for growth.

What Other States Do

Most states include mission and vision statements on their websites, as well as in their governance documents. While most of the statements focus on education data and student outcomes, many also make reference to workforce connections and outcomes. The examples below from Michigan and Minnesota provide two illustrations.

Michigan's mission and vision statements, while mentioning workforce outcomes, focus primarily on education data.[3]

> ***Our Vision:*** *Become the single source for the most comprehensive, accurate, and useful information about the performance of Michigan's public schools and students.*
>
> ***Our Mission:*** *The mission of the Center for Educational Performance and Information is to coordinate collections, connections, and reporting of education data in Michigan. CEPI was established by the Michigan Legislature (Sect. 94a of the State School Aid Act) to coordinate the collection, management, and reporting of education data, from early childhood through K–12 and postsecondary education, and into the workforce. CEPI facilitates efficient data gathering to reduce the administrative burden on reporting entities while ensuring student privacy.*
>
> *Not only does CEPI help schools comply with federal and state reporting requirements, its sophisticated longitudinal data system connects the data in powerful ways to help parents, educators, and policymakers make decisions that can improve student outcomes. One way CEPI delivers Michigan education information to stakeholders is through MI School Data.*

[3] Learn more about the Michigan system at https://www.michigan.gov/cepi/0,4546,7-113-985_71769---,00.html

A small number of states include connections to health and social services data. Minnesota has built two data systems—one focused on early childhood and one on longitudinal education data—that span education, health and social services, and workforce data.[4]

> *Minnesota's Early Childhood Longitudinal Data System (ECLDS) web tool combines data collected by the Department of Education, the Department of Human Services, and the Department of Health into one online, interactive database. The system shows population results on children's growth and achievement in relation to their participation in a variety of educational and social programs over time.*
>
> *ECLDS is the early childhood companion site to Minnesota's Statewide Longitudinal Education Data System (SLEDS). SLEDS is managed jointly by the Minnesota Office of Higher Education (OHE), the Minnesota Department of Education (MDE), and the Minnesota Department of Employment and Economic Development (DEED). Together, the two sites, ECLDS and SLEDS, form a P–20 education data system, which captures, analyzes, and uses student data from preschool to high school, college, and the workforce.*

Current Status in California

At its February 2020 meeting, the workgroup discussed possible mission and vision statements, and agreed it would like both a one-sentence vision statement as well as a longer statement with more detail. The updated statements below reflect the input received at the meeting, as well as the decisions made in April about the types of information that will be provided in phase one of the state data system.

> **Draft Vision**
>
> *The Cradle-to-Career Data System provides Californians with information to improve education, economic, and health outcomes for individuals, families, and communities.*

[4] Learn more about the Minnesota system at http://ecIds.mn.gov/#about

Draft Mission

All Californians should have the opportunity to secure an education, achieve economic mobility, and live healthy lives. To attain this goal, individuals, education providers, health and social service professionals, researchers, advocates, and policymakers should have access to information that allows them to evaluate investments of time and money. Furthermore, individuals should be supported to efficiently navigate transitions between education segments. By identifying, sharing, and strengthening the factors that help Californians meet critical milestones in the pipeline from early care to education to employment, we can work together to improve equitable outcomes for individuals, families, and communities.

The edits made to these draft statements aim to clarify that the goal of the state data system is to evaluate investments across multiple sectors, support transitions across segments, and identify factors that contribute to positive outcomes in the education pipeline. In addition, the statements now articulate that the information should be publicly available.

Considerations

To support final editing of the vision and mission statements, the workgroup needs to determine whether the state data system provides information on the education-to-work pipeline (contextualized by health, social, and economic factors) versus information on human progress toward health and economic outcomes (contextualized by education factors). This distinction is necessary because it influences the underlying data structure and legal agreements necessary to build a state data system.

Building a state data system requires careful consideration of governance, hosting, and technical architecture to ensure compliance with federal privacy laws (which will be further explored at the June workgroup meeting). It is critical to remember that compliance with these laws depends upon the purpose for which information is collected, shared, and utilized. In the education and health sectors, those purposes are generally distinct and cannot be legally comingled. Inclusion of health data requires additional layers of protections, including, but not limited to:

- For health data, information must be hosted by an entity that is covered under the Health Insurance Portability and Accountability Act (HIPAA)[5]
- For education data, the HIPAA covered entity can only receive information to evaluate a state or federal educational program[6]
- For combined data, the purpose of the analysis must be formally approved by an institutional research board (IRB) to ensure it both meets HIPAA guidelines and the FERPA condition for evaluating a state or federal education program

To overcome these obstacles, the state data system would need to store the health and education data separately and only link it for allowable purposes, as defined by federal education and health and human services agencies. Furthermore, when data are linked, they can only generate de-identified information. This means that tools such the P20W data set that populates the public dashboards and query tools, the firewalled data set for partner entities, and data produced for approved research projects would be single-point-in-time snapshots. Any subsequent changes made by partner entities to the underlying data would not be reflected unless individual records were rematched and the tools were refreshed with a new cut of the data. This would mean there could be inconsistencies between data produced by the partner entities and the state data system. It would also make it difficult for researchers to conduct analyses on the same cohorts of individuals over time.

In addition to the implications of integrating health data, at the April workgroup meeting, partner entities noted the importance of grounding the operational tools that are being considered for phase one—the California College Guidance Initiative and eTranscript California—within the goals of the state data system. These goals could be spelled out more specifically in the mission statement.

Finally, in responses to the community survey, some respondents noted the importance of providing students with mechanisms for exploring career options, both as high school students and later in life, due to the economic impact of the pandemic. If prioritized by the workgroup, this concept could also be addressed in the mission statement.

[5] Learn more about health data at https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html

[6] Learn more about sharing education data in integrated data systems at https://studentprivacy.ed.gov/sites/default/files/resource_document/file/IDS-Final_0.pdf

Next Steps

Action

- Finalize the vision and mission statements

Questions

- Should the explicit mention of health be kept, and should the focus on health extend beyond the early care context?
- Does the revised statement capture how the broader mission is served by operational tools to support education planning, electronic transcripts, and eligibility for supports?
- Should there be a stronger emphasis on employment planning in the operational tools, given the economic impacts of the pandemic?

Process

- The vision and mission statements will be discussed at the May workgroup meeting, a small group will prepare proposed wording, and partner entities will vote on final statements at the June workgroup meeting

Purpose and Scope

In a governance framework, the purpose and scope of the state data system is often recorded in a research agenda. Identifying the key questions that should be answered first can help to focus partners on the data elements that should be included in the system. The purpose and scope statement also clarifies what can be done with the information once it is linked.

What Other States Do

Most states provide the general framework of their research agenda on their websites. While some states outline a fairly general research agenda, others use a more specific set of questions, organized by category. For example, Maryland's research agenda is focused on critical education and workforce transition points and outcomes, and specifically, what happens to individuals both before and after those critical transitions.

The research agenda includes a set of 21 questions organized in four categories. See the box below for sample questions from each category.

Maryland's Research Agenda: Sample Questions[7]

- K–12 Readiness
  - What is the impact of early childhood education experiences and programs on children's school readiness and K–12 outcomes?
- Postsecondary Readiness and Access
  - What percentage of Maryland high school exiters entering college are assessed to need to take developmental courses and in what content areas?
  - Which financial aid programs are most effective in improving access and success (i.e., retention and graduation) for Maryland students?
- Postsecondary Completion
  - How likely are students placed in developmental courses to persist in postsecondary education and transfer and/or graduate?
  - What are the differences in performance, retention, and graduation, including time to degree, of students who initially matriculate at a Maryland community college and transfer to a Maryland four-year institution versus those who initially matriculate at a Maryland four-year?
- Workforce Outcomes
  - What are the educational and labor market outcomes for individuals who use federal and state resources to obtain training at community colleges or other postsecondary institutions?
  - What are the workforce outcomes for Maryland students who earn a high school diploma (via high school graduation or GED) but do not transition to postsecondary education or training?

The purpose and scope statement for state data systems also specify that they only provide data that are linked across more than one entity or sector. For example, Washington's Educational Research and Data Center notes that it "limits itself to

[7] Read the complete Maryland Research Agenda at https://mldscenter.maryland.gov/ResearchAgenda.html

seeking answers to cross-sector research questions (questions that cannot be answered by institutions within one sector)."[8] In Minnesota, the state data system lists as a first principle that the system will focus on providing cross-sector, linked data and analysis. The governance policy refers any data requests "limited to just one sector (K–12 education, postsecondary education, or employment) to the appropriate agency or partner data provider."[9]

Uses for California's State Data System[10]

The legislation suggests that the state data system should serve students and families by doing all of the following:

- Identifying and tracking predictive indicators to enable parents, teachers, health and human services providers, and policymakers to provide appropriate interventions and supports to address disparities in opportunities and improve outcomes for all students
- Creating direct support tools for teachers, parents, advisors, and students
- Enabling agencies to plan for and optimize educational, workforce, and health and human services programs
- Advancing academic and governmental research on improving policies from birth through career

Current Status in California

In addition to determining the research questions that will drive the data included in the first phase of the state data system, the partner entities are tasked with determining how this information would be made available to people in a broad range of roles. These decisions are guided by language in the Act, which outlines the purpose of the state data system. First, the legislation highlights six priority policy topics, which are currently being examined by the Research Agenda Subcommittee (subcommittee). The subcommittee is generating a list of questions for each topic and specifying which

[8] Learn more about the Washington system at https://erdc.wa.gov/about-us-0

[9] Read the Minnesota's SLEDS Data Access Management Policy at http://sleds.mn.gov/#research

[10] Read the relevant text from the Act at https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=EDC&division=1.&title=1.&part=7.&chapter=8.5.&article=

questions would be answered through dashboards, could be investigated through query tools, or should inform the prioritization of requests to conduct research studies. The legislation also articulates how this information should be used (see box on previous page).

Several decisions have already been made regarding how information from the state data system will be accessed, including identifying specific analytical and operational tools.

Analytical Tools

- A centralized data set that provides information to the public through a multi-level dashboard and a query builder
- A data request process that allows authorized researchers to use deidentified, unitary data to conduct more in-depth analyses

Operational Tools

- By scaling the California College Guidance Initiative (CCGI), provide tools and curriculum for college and career guidance to K–12 students, allow families and educators to monitor factors that influence college-going for individual students, streamline the application process for college and financial aid, and support the cleanup of underlying data used for college admissions and financial aid
- By building upon the eTranscript California model, streamline the movement of transcripts across educational segments and expand records such as competency-based credentials, which are of value for adults pursuing employment-related training
- By integrating social service records into electronic transcripts, enable students to authorize information-sharing that would qualify them for support services

When evaluating the operational tools use case following the April workgroup meeting, one partner entity noted the need for more guidance related to transfer pathways for community college students. This concept could be added to a framework that evaluates how well CCGI and eTranscript California address the goals of the state data system.

In addition, some members of the advisory groups, the Research Agenda Subcommittee, and the public have expressed a desire to use the state data system to access information from a single partner entity. For example, a public dashboard might

show the impact of chronic absenteeism on high school graduation in addition to postsecondary enrollment. Or, a query builder tool could provide information on the impact of community college course taking patterns on both becoming transfer prepared and transferring. This topic has been discussed at several advisory group, subcommittee, and workgroup meetings and the use cases imply that only linked data will be provided. However, no formal decision has been made.

Considerations

The policy priorities outlined in the legislation focus on the education-to-workforce pipeline. While the questions imply a focus on linked data, they do not include specific questions regarding health, social service, teacher training, institutional, or economic contexts. If the mission includes this broader scope, then the research agenda may need to be expanded to more specifically address questions on topics beyond education, financial aid, and employment.

Next Steps

Action

- Finalize a decision about whether the system only releases intersegmental data
- Determine the scope of the research agenda

Questions

- Should the state data system only provide information that links data from two or more providers, and refer requests for information from a single agency back to that partner entity?
- Should the research agenda be expanded to specifically address the impact of health and social services investments?
- Should additional policy topics be added to the research agenda in light of the pandemic?
- Should the operational tools include guidance on transfer pathways?

Process

- The question of access to data from an individual partner will be voted on at the May workgroup meeting

- The question of whether to expand the research agenda will be discussed at the May workgroup meeting and then delegated to the Research Agenda Subcommittee. A draft research agenda will be shared with the workgroup in September.

Decision-Making Model

In addition to establishing purpose and scope, the governance framework spells out how decisions will be made.

What Other States Do

States create documents that codify how decisions are made, such as memoranda of understanding, bylaws, or charters. According to the Data Quality Campaign, states that put their committee and decision-making models into legislation are best able to sustain functional intersegmental systems.[11]

Most states use a committee structure, with clearly outlined roles and responsibilities. Many create a governing board that has decision-making authority, with separate committees to address research, technical, and communication issues. Often, the governing board includes both partner entities and external parties that represent different stakeholder voices. Some states also include an executive committee comprised of the system heads of all partner entities.

Maryland's governing board, for example, is made up of 12 members, including seven designated by statute and five appointed by the governor with the advice and consent of the state senate. Those designated by statute include representatives of the entities that contribute data to the system, including the university system chancellor, state school superintendent, higher education secretary, labor secretary, Morgan State University president, Maryland Association of Community Colleges executive director, and Maryland Independent Colleges and Universities Association president. Among the governor's appointees, one must be a representative of local superintendents of schools and another must have expertise in large data systems and data security. The chair of the governing board is appointed by the governor. To ensure that system

[11] Personal communication with Paige Kowalski, Data Quality Campaign, May 7, 2020.

management and information output meets the needs of various stakeholders, the state also convenes two committees: one on data and one on research and policy.[12]

In Virginia, the governing board is made up of representatives from the partner agencies and is responsible for setting the vision for the state data system and establishing policies and procedures. In addition, there are four subcommittees.[13]

- The technical subcommittee addresses the maintenance and operation of the data infrastructure, as well as privacy and security.
- The research subcommittee evaluates questions submitted by partner entities, develops a consolidated list for approval by the governing board, oversees progress of research projects, and provides support for proper use of data.
- The communications subcommittee oversees strategic communications with the public and develops one-page summaries of research conducted using state data system information.
- The legislative subcommittee works with the state assembly on financing the state data system.

Prior Governing Board Recommendations

The Education Data and Information Act of 2008 (SB 1298, Chapter 561) authorized a planning process similar to the one being undertaken for the Cradle-to-Career data system. In its recommendations, the governance workgroup proposed that the majority of governing board members should come from outside of the partner entities, including bipartisan representation from the legislature and administration and representatives from education institutions.

Current Status in California

The Act empowers the workgroup to "advise ongoing efforts to develop, administer, and enhance the data system." These responsibilities could be extended once the

[12] Learn more about Maryland's system at https://mldscenter.maryland.gov/Governance.html

[13] Learn more about the Virginia system at http://vlds.virginia.gov/media/1087/vlds_book_of_dg.pdf

state data system is in place, to create a governing board with decision-making authority.

As part of the planning process for the state data system, five subcommittees have been developing recommendations for the workgroup on the following topics: Common Identifier, Definitions, Legal, Research Agenda, and Technology & Security. The subcommittees include representatives from the partner entities, practitioners, researchers, and data system experts. The subcommittee structure parallels committees found in other states and could be established as standing committees once the state data system is developed.

In addition, two advisory groups made up of advocacy organizations, research entities, and practitioners have been tasked with ensuring that the planning process addresses the vision articulated in the Act. The Policy & Analytics Advisory Group reviews whether plans will support research, evaluation, accountability, and optimization of publicly funded services at the state level. The Practice & Operations Advisory Group examines whether plans would support improvement efforts at the institutional and regional level, a case management approach to service delivery, and tools that would be useful to students, families, and teachers.

Who Makes Decisions?

In the current planning process, the Research Agenda and Data Definitions Subcommittees are helping to determine which data elements will initially be available to the public through dashboards, query builders, or data requests. However, once the state data system is underway, a specific entity needs to be responsible for evaluating additional data elements for inclusion and clarifying the contexts in which those data points should be shared. This responsibility could be assigned to a data committee, delegated to individual partner entities, or evaluated by staff at the entity managing the state data system.

Considerations

In addition to determining whether the current workgroup, subcommittee, and advisory group structure should be maintained, the workgroup needs to identify the specific membership, roles, and responsibilities of each committee. The workgroup should also evaluate whether legislation should be passed that codifies the committee structure.

Next Steps

Action

- Develop committee structure, responsibilities, and decision-making process

Questions

- Should any of the current subcommittees be continued as part of the governance framework after the planning phase? Which ones? What responsibilities would they be given?
- Should additional committees be created, such as an executive committee made up of system heads, a communications committee, a legislative committee, or a committee specific to the operational tools (CCGI and eTranscript California)?
- Should the current advisory groups continue to provide input or should some members of the advisory groups be added to the governing board? What other voices should be added to the governing board or other committees? How should those members be chosen?
- How will decisions be made within the committees?
- Should legislation be passed that codifies the committee structure?

Process

- This topic will be discussed at the May workgroup meeting, followed by a small group to develop a proposal, and a vote at the June workgroup meeting